Modelling of Sanitary Hot Water Energy Consumption Using Adaptive Neuro-Fuzzy Inference Systems

by

George William Blignault

Thesis presented in partial fulfilment of the requirements for the degree of Master of Engineering in the Faculty of Engineering at Stellenbosch University

Study leader: Prof. HJ Vermeulen

December 2016

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 2016

Acknowledgements

I would like to thank the following people for their contribution to this project:

- Prof HJ Vermeulen for his expert and relentless guidance throughout the course of this project.
- My parents Ernie and Suzette Blignault for years of love and support.
- Vicki, Tielman, Almero, Gunnar, Sven, Edwin, and Chris. For the many good times in E222 and beyond.
- Ryno Lochner for his assistance with the access and acquisition of the US heat pump monitoring data.
- South African Weather Service for providing weather data records vital to the research, and for their quick responses to queries concerning said data.

Abstract

The introduction of Energy Management (EM) interventions aimed at reducing electrical energy consumption in the residential, commercial and industrial load sectors has expanded rapidly on a global scale in recent years. These programs are driven by environmental concerns, capacity constraints experienced in generation and transmission and the need to improve end-use efficiency. The implementation of EM schemes often involves financial incentives funded by governments and utilities. Measurement and Verification (M&V) performance assessments aimed at determining the savings impacts form an integral part of the management of these EM incentive programmes. M&V baseline development involves the development and implementation of accurate models that relate the energy consumption of a targeted load to variable energy-governing factors in order to determine actual savings impacts.

The electrical energy consumption associated with sanitary water heating represents a sizeable component of the cumulative energy consumption associated with a number of load categories found in the various load sectors. In general, the electricity consumption profiles associated with sanitary hot water consumption correlates closely with the household electricity consumption profiles found in the residential load sector, particularly in the sense that it is influenced by the same socio-economic factors and human behavioural patterns. Soft computing methods have been employed successfully for residential load prediction, as these are tolerant of stochastic behaviour and uncertainty and do not require exact input to output matching. Particular success in the field of residential Short Term Load Forecasting (STLF) has been achieved using Adaptive Neuro-Fuzzy Inference Systems (ANFIS).

An ANFIS load forecasting model with a long prediction horizon of up to a year is found to be capable of reasonable modelling accuracy for the estimation of the time-series profile of a system. It also exhibits very good prediction accuracy when calculating the total energy use over time of that profile. The load data used in this study is of student residence heat pump power consumption profiles and spans over four years with 48 samples for each day. The training inputs that are considered other than the load are the time of the day, the day of the week, the day of the year and the temperature.

After the proof of concept, a comparative case study is performed with the view to explore optimal configurations of differing inputs to the ANFIS method. The effects of compartmentalising the dataset into subsets representing different characteristics, thereby

deriving different models representing different cyclic periods, are also explored. It is found that compartmentalising the load model into 48 ANFIS sub-models, each serving a specific half-hourly time period in the day, results in the most best modelling accuracy.

Opsomming

Die bekendstelling van Energie Bestuur (EB) ingrypings wat gerig is op die vermindering van elektriese energieverbruik in die residensiële, kommersiële en industriële lassektore het vinnig uitgebrei op 'n globale skaal in onlangse jare. Hierdie programme word gedryf deur omgewinskwessies, ervaring van kapasiteitsbeperkings tydens die opwekking en transmissie van krag en die behoefte om einde gebruik doeltreffendheid te verbeter. Die implementering van EB skemas behels dikwels finansiële aansporings wat deur regerings en kragopwekkende instansies befonds word. Meting en verifikasie (M&V) prestasiebeoordelings wat gemik is op die bepaling van die besparings impak vorm 'n integrale deel van die bestuur van hierdie EB aansporings programme. M&V basislyn ontwikkeling behels die ontwikkeling en implimentering van akkurate modelle wat die energieverbruik van n teikenlas aan energie veranderlike faktore bind ten einde die besparingsimakte te bepaal.

Die elektriese energieverbruik wat verband hou met sanitêre water verwarming verteenwoordig 'n groot komponent van die kumulatiewe energieverbruik wat verband hou met 'n aantal las kategorieë in die verskeie lassektore. Die elektrisiteitsverbruik profiele wat verband hou met sanitêre warm waterverbruik korreleer oor die algemeen met die huishoudelike elektrisiteitsverbruik profiele van die residensiële sektor, veral in die sin dat dit beïnvloed word deur dieselfde sosio-ekonomiese faktore en menslike gedragspatrone. Sagte rekenaar metodes is al suksesvol gemplimenteer vir residensiële lasvoorspelling, aangesien dit verdraagsaam van stogastiese gedrag en onsekerheid is en nie dat die insette en uitsette presies ooreen stem nie. Sukses is veral in die gebied van residensiële Kort Termyn lasvooruitskatting (KTLV) behaal met behulp van Adaptive Neuro-Fuzzy Inferensie Systems (ANFIS).

Dit is gevind dat 'n ANFIS las voorspellings model met 'n lang voorspelling horison van tot 'n jaar in staat is tot redelike modellering akkuraatheid vir die skatting van die tyd-reeks profiel van 'n stelsel. Dit vertoon ook baie goeie voorspelling akkuraatheid by die berekening van die totale energieverbruik van die profiel oor tyd. Die las data wat in hierdie studie gebruik word is van studentekoshuise se warmtepomp kragverbruik profiele en strek oor vier jaar met 48 monsters vir elke dag. Die opleiding insette wat beskou word, anders as die las, is die tyd van die dag, die dag van die week, die dag van die jaar en die temperatuur.

Na afloop van die bewys van die konsep is 'n vergelykende gevallestudie uitgevoer met die doel om optimale konfigurasies van verskillende insette tot die ANFIS metode te verken. Die

gevolge van kompartimentalisering van die datastel in sub-versamelings wat verskillende eienskappe verteenwoordig, en dus verskillende modelle verteenwoordigend van verskillende sikliese periodes aflei, word ook ondersoek. Daar word bevind dat kompartimentalisering van die lasmodel in 48 ANFIS sub-modelle, wat elk 'n spesifieke half-uurlikse tydperk in die dag bedien, lei tot die beste modelleringsakkuraatheid.

Table of Contents

Chapter 1: Project Motivation and Project Description 1

1.1 Introduction 1

1.2 Project motivation 3

1.2.1 Overview 3

1.2.2 Medium-term load prediction 3

1.2.3 Measurement and Verification performance evaluation of energy management interventions 4

1.2.4 Potential of Adaptive Neuro-Fuzzy Inference Systems for water heating energy consumption modelling 5

1.3 Project description 6

1.3.1 Overview 6

1.3.2 Research objectives 6

1.3.3 Key questions 7

1.3.4 Research Tasks 8

1.4 Research methodology 9

1.5 Overview of report 10

Chapter 2: Literature Review 12

2.1 Overview 12

2.2 Sanitary hot water heating 12

2.2.1 Overview 12

2.2.2 Diurnal cycles in sanitary hot water use in different load sectors 12

2.2.3 Weekly cycles in sanitary hot water use 13

2.2.4 Seasonal effects on average heat pump power consumption 13

2.3 Theory of machine learning elements 14

2.3.1 Introduction and scope 14

2.3.2 General principles....14
2.3.3 Inductive and deductive learning....16
2.3.4 Fuzzy systems....17
2.3.5 Adaptive networks....20
2.4 ANFIS theoretical framework....21
2.4.1 ANFIS model architecture....21
2.4.2 Hybrid learning rule....23
2.5 ANFIS implementation in the MATLAB environment....23
2.5.1 The MATLAB platform....23
2.5.2 Fuzzy Logic Toolbox....24
2.5.3 ANFIS functionality....24
2.6 Prediction accuracy metrics....25
2.6.1 Overview....25
2.6.2 Error....25
2.6.3 RMSE....26
2.6.4 MAE....26
2.6.5 MAPE....26
2.7 Forecasting with ANFIS....27
Chapter 3: Model Implementation....28
3.1 Overview....28
3.2 Heat pump load profiles....28
3.3 Ambient temperature profiles....31
3.4 Hypothesis formulation....33
3.5 Data conditioning....35
3.5.1 Equal dimensionality formatting....35
3.5.2 Appropriate variable scaling....35

3.5.3 Outlier removal 35
3.5.4 Final training data set 40
3.6 Statistical analysis of training data 41
3.6.1 Overview 41
3.6.2 Heat pump load data 41
3.6.3 Temperature data 43
3.7 Experiment design 44
3.7.1 Overview 44
3.7.2 Experimental model configurations 45
3.7.3 Partitioning of training data 46
3.7.4 K-fold cross validation 46
3.7.5 Cross validation results 47
3.7.6 Design choices 51
Chapter 4: Software Implementation 53
4.1 Overview 53
4.2 Data formatting 53
4.2.1 Formatting objective 53
4.2.2 Formatting load data 53
4.2.3 Formatting of temperature data 54
4.2.4 Data sanitation 55
4.2.5 Load data pre-processing 55
4.3 Training and testing module 55
4.3.1 Data flow 55
4.3.2 Dataset input constructor operations 56
4.3.3 Training and experimental ANFIS configurations 57
4.4 Accuracy report module 59

Chapter 5: Modelling Results and Performance Evaluation 61
5.1 Overview 61
5.2 Case study objectives 61
5.3 Case study methodology 61
5.3.1 Overview 61
5.3.2 Case study parameter summary 62
5.3.3 Graph model outputs 62
5.3.4 Performance evaluation 63
5.3.5 Test for over-fitting 63
5.4 Group A case study 64
5.4.1 Model A1 64
5.4.2 Model A2 68
5.4.3 Model A3 70
5.4.4 Model A4 73
5.5 Group B case study 76
5.5.1 Model B1 76
5.5.2 Model B2 79
5.5.3 Model B3 82
5.6 Group C case study 84
5.6.1 Model C1 84
5.6.2 Model C2 87
5.6.3 Model C3 90
5.7 Group D case study 92
5.7.1 Model D1 92
5.7.2 Model D2 95
5.7.3 Model D3 97

5.7.4 Model D4 100
5.8 Summary of case-study results 102
5.9 Computational complexity 103
5.10 Analysis of time-series prediction accuracies 105
5.10.1 Overview 105
5.10.2 Accuracies of timeslots 105
5.10.3 Accuracies of weekday categories 106
5.11 Analysis of total energy prediction accuracies 106
Chapter 6: Conclusions and Recommendations 109
6.1 Conclusions 109
6.1.1 Overview 109
6.1.2 Application of ANFIS to model and predict the hot water energy consumption profiles associated with student residences 109
6.1.3 Comparative case study of different ANFIS implementations 112
6.1.4 Feasibility for practical applications 114
6.2 Recommendations for future work 115
Appendix A Model Outputs for Cross Validation Training 119
Appendix A.1 Model A1 119
Appendix A.2 Model A2 122
Appendix A.3 Model A3 125
Appendix A.4 Model A4 128
Appendix A.5 Model B1 131
Appendix A.6 Model B2 134
Appendix A.7 Model B3 137
Appendix A.8 Model C1 140
Appendix A.9 Model C2 143
Appendix A.10 Model C3 146

Appendix A.11 Model D1 149
Appendix A.12 Model D2 152
Appendix A.13 Model D3 155
Appendix A.14 Model D4 158
Appendix B Error Percentages 161
Appendix B.1 Scatterplot of the percentage errors over observed load size 161

List of figures

Figure 1-1: Diagram illustrating the facets of the motivations that are to drive the research. 3
Figure 2-1: Partitioning of training and validation data in the case of inductive learning [18]. 16
Figure 2-2: Example partitioning of training and validation data in the case of deductive learning[18]. 17
Figure 2-3: Boiling temperature as a function of water temperature membership: (a) A crisp value, (b) A crisp set and (c) A fuzzy set [18]. 18
Figure 2-4: Functional topology of a fuzzy inference system [9]. 20
Figure 2-5: Example of a feed forward adaptive network [25]. 21
Figure 2-6: ANFIS architecture for a two input, one output inference system[26]. 23
Figure 3-1: Combined heat pump load profile for 2009. 29
Figure 3-2: 3D plot of the combined heat pump half-hourly averaged power profile for the calendar years 2009 to 2011. 31
Figure 3-3: Heat map plot of the combined heat pump half-hourly averaged power profile for the calendar years 2009 to 2011. 31
Figure 3-4: Temperature as recorded at the CTIA weather station for 2009. 32
Figure 3-5: 3D plot of the temperature as measured at the CTIA weather profile for the calendar years 2009 to 2011. 33
Figure 3-6: 2D heat map plot of the temperature as measured at the CTIA weather station profile for the calendar years 2009 to 2011. 33
Figure 3-7: Heat map of the combined heat pump load profile the academic reses periods removed. 37
Figure 3-8: Heat map of the combined heat pump load profile with academic reses periods, public holidays and second examination periods removed. 38
Figure 3-9: Heat map of the combined heat pump average power profile for 2008. 39
Figure 3-10: Heat map of the half-hourly average power profile of the combined heat pump load after data conditioning. 40
Figure 3-11: Heat map of the half-hourly averaged temperature profile after data conditioning. 41

Figure 3-12: Histograms of half-hourly energy consumption for the calendar years from 2008 to 2011.....43
Figure 3-13: Histograms of daily energy consumption for the calendar years from 2008 to 2011.....43
Figure 3-14: Histograms of the average ambient temperature for the calendar years from 2008 to 2011.....44
Figure 3-15: Heat map of the normalised combined heat pump power profile for 2008.....48
Figure 3-16: 3D plot of the normalised combined heat pump power profile for 2008.....48
Figure 3-17: Heat map of the normalised combined heat pump power profile for 2009.....49
Figure 3-18: 3D plot of the normalised combined heat pump power profile for 2009.....49
Figure 3-19: 2D plot of the normalised combined heat pump power profile for 2010.....50
Figure 3-20: 3D plot of the normalised combined heat pump power profile for 2010.....50
Figure 3-21: Heat map of the normalised combined heat pump power profile for 2011.....51
Figure 3-22: 3D plot of the normalised combined heat pump power profile for 2011.....51
Figure 4-1: Data flow and operations for the formatting of the heat pump load data.....54
Figure 4-2: Data flow and operations for the formatting of the temperature data.....55
Figure 4-3: Data flow for model training and validation.....56
Figure 4-4: Operations of input constructor.....57
Figure 4-5: Single ANFIS training configuration.....57
Figure 4-6: Weekday category sub-model ANFIS configuration and operations.....58
Figure 4-7: Time of day sub-model ANFIS configuration and operations.....58
Figure 4-8: Time of day in weekday category sub-model ANFIS configuration and operations.....59
Figure 5-1: Training and checking error RMSE values over course of training epochs for each checking year in the k-fold cross validation method.....64
Figure 5-2: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model A1.....65
Figure 5-3: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model A1.....66

Figure 5-4: MAPE over each cross validation fold over each daily timeslot using model A1. 67
Figure 5-5: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model A2 68
Figure 5-6: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model A2 69
Figure 5-7: MAPE over each cross validation fold over each daily timeslot using model A2. 70
Figure 5-8: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model A3 71
Figure 5-9: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model A3 72
Figure 5-10: MAPE over each cross validation fold over each daily timeslot using model A3. 73
Figure 5-11: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model A4 74
Figure 5-12: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model A4 75
Figure 5-13: MAPE over each cross validation fold over each daily timeslot using model A4. 76
Figure 5-14: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model B1 77
Figure 5-15: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model B1 77
Figure 5-16: MAPE over each cross validation fold over each daily timeslot using model B1. 78
Figure 5-17: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model B2 80
Figure 5-18: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model B2 80
Figure 5-19: MAPE over each cross validation fold over each daily timeslot using model B2. 81

Figure 5-20: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model B3.....82
Figure 5-21: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model B3.....83
Figure 5-22: MAPE over each cross validation fold over each daily timeslot using model B3. 84
Figure 5-23: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model C1.....85
Figure 5-24: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model C1.....86
Figure 5-25: MAPE over each cross validation fold over each daily timeslot using model C1.....87
Figure 5-26: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model C2.....88
Figure 5-27: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model C2.....88
Figure 5-28: MAPE over each cross validation fold over each daily timeslot using model C2.....89
Figure 5-29: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model C3.....90
Figure 5-30: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model C3.....91
Figure 5-31: MAPE over each cross validation fold over each daily timeslot using model C3.....92
Figure 5-32: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model D1. 93
Figure 5-33: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model D1. 93
Figure 5-34: MAPE over each cross validation fold over each daily timeslot using model D1. 94
Figure 5-35: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model D2. 95

Figure 5-36: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model D2. 96
Figure 5-37: MAPE over each cross validation fold over each daily timeslot using model D2. 97
Figure 5-38: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model D3. 98
Figure 5-39: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model D3. 98
Figure 5-40: MAPE over each cross validation fold over each daily timeslot using model D3. 99
Figure 5-41: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model D4. 100
Figure 5-42: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model D4. 101
Figure 5-43: MAPE over each cross validation fold over each daily timeslot using model D4. 102
Figure 5-44: Percentage error of model A1 predictions over size of observed load at error. 106
Figure A-1: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model A1 119
Figure A-2: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model A1 119
Figure A-3: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model A1 120
Figure A-4: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model A1 120
Figure A-5: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model A1 121
Figure A-6: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model A1 121
Figure A-7: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model A2 122

Figure A-8: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model A2 122
Figure A-9: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model A2 123
Figure A-10: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model A2 123
Figure A-11: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model A2 124
Figure A-12: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model A2 124
Figure A-13: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model A3 125
Figure A-14: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model A3 125
Figure A-15: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model A3 126
Figure A-16: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model A3 126
Figure A-17: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model A3 127
Figure A-18: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model A3 127
Figure A-19: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model A4 128
Figure A-20: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model A4 128
Figure A-21: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model A4 129
Figure A-22: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model A4 129
Figure A-23: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model A4 130

Figure A-24: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model A4 130
Figure A-25: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model B1 131
Figure A-26: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model B1 131
Figure A-27: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model B1 132
Figure A-28: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model B1 132
Figure A-29: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model B1 133
Figure A-30: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model B1 133
Figure A-31: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model B2 134
Figure A-32: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model B2 134
Figure A-33: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model B2 135
Figure A-34: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model B2 135
Figure A-35: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model B2 136
Figure A-36: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model B2 136
Figure A-37: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model B3 137
Figure A-38: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model B3 137
Figure A-39: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model B3 138

Figure A-40: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model B3. 138
Figure A-41: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model B3. 139
Figure A-42: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model B3. 139
Figure A-43: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model C1. 140
Figure A-44: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model C1. 140
Figure A-45: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model C1. 141
Figure A-46: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model C1. 141
Figure A-47: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model C1. 142
Figure A-48: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model C1. 142
Figure A-49: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model C2. 143
Figure A-50: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model C2. 143
Figure A-51: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model C2. 144
Figure A-52: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model C2. 144
Figure A-53: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model C2. 145
Figure A-54: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model C2. 145
Figure A-55: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model C3. 146

Figure A-56: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model C3. 146
Figure A-57: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model C3. 147
Figure A-58: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model C3. 147
Figure A-59: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model C3. 148
Figure A-60: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model C3. 148
Figure A-61: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model D1. 149
Figure A-62: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model D1. 149
Figure A-63: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model D1. 150
Figure A-64: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model D1. 150
Figure A-65: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model D1. 151
Figure A-66: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model D1. 151
Figure A-67: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model D2. 152
Figure A-68: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model D2. 152
Figure A-69: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model D2. 153
Figure A-70: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model D2. 153
Figure A-71: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model D2. 154

Figure A-72: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model D2. 154
Figure A-73: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model D3. 155
Figure A-74: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model D3. 155
Figure A-75: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model D3. 156
Figure A-76: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model D3. 156
Figure A-77: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model D3. 157
Figure A-78: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model D3. 157
Figure A-79: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model D4. 158
Figure A-80: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model D4. 158
Figure A-81: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model D4. 159
Figure A-82: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model D4. 159
Figure A-83: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model D4. 160
Figure A-84: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model D4. 160
Figure B-1: Percentage error of model A2 predictions over size of observed load at error. 161
Figure B-2: Percentage error of model A3 predictions over size of observed load at error. 161
Figure B-3: Percentage error of model A4 predictions over size of observed load at error. 162

Figure B-4: Percentage error of model B1 predictions over size of observed load at error. 162
Figure B-5: Percentage error of model B2 predictions over size of observed load at error. 163
Figure B-6: Percentage error of model B3 predictions over size of observed load at error. 163
Figure B-7: Percentage error of model C1 predictions over size of observed load at error. 164
Figure B-8: Percentage error of model C2 predictions over size of observed load at error. 164
Figure B-9: Percentage error of model C3 predictions over size of observed load at error. 165
Figure B-10: Percentage error of model D1 predictions over size of observed load at error. 165
Figure B-11: Percentage error of model D2 predictions over size of observed load at error. 166
Figure B-12: Percentage error of model D3 predictions over size of observed load at error. 166
Figure B-13: Percentage error of model D4 predictions over size of observed load at error. 167

List of Tables

Table 1-1: Guide to which sections of the document specifically relate to steps of the chosen research methodology. 10
Table 3-1: Summary of heat pumps and residences. 28
Table 3-2: Documented South-African public holiday dates throughout the load data range, underlined dates indicate unobserved public holidays. 36
Table 3-3: Documented start and end dates for classes, first exam periods, and second exam periods of Stellenbosch University over the load data range. 36
Table 3-4: Dates removed due to abnormal behaviour in transitional periods 39
Table 3-5: Mean and standard deviation statistics for the half-hourly and daily energy consumption for the calendar years from 2008 to 2011 42
Table 3-6: Mean temperature and standard deviation of temperatures for the calendar years from 2008 to 2011 44
Table 3-7: K-fold cross validation training and testing data allocation for each fold. 47
Table 5-1: Model configurations for case studies. 62
Table 5-2: MAPE of each cross validation fold over each weekday category using model A1. 66
Table 5-3: Percentage error of model A1 in estimating yearly energy consumption within the training intervals. 67
Table 5-4: MAPE of each cross validation fold over each weekday category using model A2. 69
Table 5-5: Percentage error of model A2 in estimating yearly energy consumption within the training intervals. 70
Table 5-6: MAPE of each cross validation fold over each weekday category using model A3. 72
Table 5-7: Percentage error of model A3 in estimating yearly energy consumption within the training intervals. 73
Table 5-8: MAPE of each cross validation fold over each weekday category using model A4. 75
Table 5-9: Percentage error of model A4 in estimating yearly energy consumption within the training intervals. 76

Table 5-10: MAPE of each cross validation fold over each weekday category using model B1. 78
Table 5-11: Percentage error of model B1 in estimating yearly energy consumption within the training intervals. 79
Table 5-12: MAPE of each cross validation fold over each weekday category using model B2. 81
Table 5-13: Percentage error of model B2 in estimating yearly energy consumption within the training intervals. 81
Table 5-14: MAPE of each cross validation fold over each weekday category using model B3. 83
Table 5-15: Percentage error of model B3 in estimating yearly energy consumption within the training intervals. 84
Table 5-16: MAPE of each cross validation fold over each weekday category using model C1. 86
Table 5-17: Percentage error of model C1 in estimating yearly energy consumption within the training intervals. 87
Table 5-18: MAPE of each cross validation fold over each weekday category using model C2. 89
Table 5-19: Percentage error of model C1 in estimating yearly energy consumption within the training intervals. 89
Table 5-20: MAPE of each cross validation fold over each weekday category using model C3. 91
Table 5-21: Percentage error of model C3 in estimating yearly energy consumption within the training intervals. 92
Table 5-22: MAPE of each cross validation fold over each weekday category using model D1. 94
Table 5-23: Percentage error of model D1 in estimating yearly energy consumption within the training intervals. 94
Table 5-24: MAPE of each cross validation fold over each weekday category using model D2. 96
Table 5-25: Percentage error of model D2 in estimating yearly energy consumption within the training intervals. 97

Table 5-26: MAPE of each cross validation fold over each weekday category using model D3. 99
Table 5-27: Percentage error of model D3 in estimating yearly energy consumption within the training intervals. 99
Table 5-28: MAPE of each cross validation fold over each weekday category using model D4. 101
Table 5-29: Percentage error of model D4 in estimating yearly energy consumption within the training intervals. 102
Table 5-30: average MAPE of experimental load model cross validation. 102
Table 5-31: Duration of training of each cross validation fold for each model. 104
Table 5-32: Mean and standard deviations of the percentage error for predictions of total energy use in each year, shown for each individual group as well as that of all models. 107

Nomenclature

EM	Energy Management
EE	Energy Efficient
M&V	Measurement and verification
EEDSM	Energy efficiency and demand side management
SLA	Service level adjustments
ANN	Artificial neural network
FIS	Fuzzy inference system
ANFIS	Adaptive neuro-fuzzy inference system
MF	Membership function
PE	Percentage Error
RMSE	Root Mean Squared Error
MAE	Mean Average Error
MAPE	Mean Absolute Percentage Error
STD	Standard Deviation

Chapter 1: Project Motivation and Project Description

1.1 Introduction

The introduction of Energy Management (EM) interventions aimed at reducing electrical energy consumption in the residential, commercial and industrial load sectors has expanded rapidly on a global scale in recent years [1]. These interventions focus on strategies such as the identification and elimination of wasteful energy consumption and the introduction of modern Energy Efficient (EE) technologies to improve end-use efficiency. From an environmental perspective, EM programs are driven by the urgent need to reduce the global use of fossil fuels for the production of electrical energy [2]. In many developing economies, however, EM programmes are driven by capacity constraints experienced in the generation and transmission of electrical energy. In these cases, the EM interventions include strategies such as load-shifting with the view to reduce consumption during peak periods, thereby avoiding the need for load-shedding to maintain system stability [3].

The implementation of EM schemes often involves financial incentives funded by governments and utilities. Measurement and Verification (M&V) performance assessments aimed at determining the savings impacts, both with reference to energy savings and demand reduction, form an integral part of the management of these EM incentive programmes. Measurement and Verification baseline development, baseline adjustment and performance assessment involves the development and implementation of accurate models that relate the energy consumption of a targeted load to variable energy-governing factors in order to determine actual savings impacts, or in the case of scoping studies, to quantify the potential for savings [4].

The electrical energy consumption associated with sanitary water heating represents a sizeable component of the cumulative energy consumption associated with a number of load categories found in the various load sectors. These load categories include households in the residential sector, the hospitality and health industries in the commercial sector, labour accommodation in the industrial sector and student residences at Higher Education Institutions (HEIs) in the public sector [5][6][7]. The hot water consumption profiles for the various load categories and load sectors vary widely, depending on the nature of the economic activity and human occupation involved. The electrical energy consumption profiles associated with sanitary water are, furthermore, heavily influenced by behavioural factors

associated with socio-economic considerations such as Living Standard Measure (LSM) [8], lifestyle, cultural norms and social norms [5]. Location dependent physical factors associated with atmospheric and climatic conditions, such as ambient temperature and daylight hours, add further complexity [6] [9].

Despite the stochastic nature introduced by the above factors, the consumption profiles associated with sanitary hot water usage typically exhibit pronounced diurnal, weekly and seasonal cycles. The typical daily cycle reflects the daily human activity profile, characterised by high sanitary hot water consumption in the morning and evening periods before and after the day-time activity cycle [5]. The weekly cycle generally exhibits distinctly different daily cycles for weekdays, Saturdays and Sundays, reflecting more pronounced morning and evening peaks for weekdays as a result of the well-defined activity profiles associated with weekdays. The seasonal cycle arises from changes in ambient temperature and daylight hours and the consumption of hot water generally increases in the cold season.

In general, the electricity consumption profiles associated with sanitary hot water consumption correlates closely with the household electricity consumption profiles found in the residential load sector, particularly in the sense that it is influenced by the same socio-economic factors and human behavioural patterns. Soft computing methods have been employed successfully for residential load prediction, as these are tolerant of stochastic behaviour and uncertainty and do not require exact input to output matching. This makes it ideal for residential load modelling, where the correlation between power consumption and the independent driving factors may be vague and nonlinear. The Adaptive Neuro-Fuzzy Inference System (ANFIS) [10] learning method, which is a hybrid of Fuzzy Inference Systems (FIS) and back propagated Artificial Neural Networks (ANN), has been employed successfully for residential Short-Term Load Forecasting (STLF), i.e. hour ahead forecasting, for after diversity demand modelling of residential networks [11], [12]. These short-term models exhibit Mean Average Percentage Errors (MAPE) as low as 2%. The models are typically trained and tested using data for stable summer and winter months, because of the transitional nature of the load profiles associated with other seasons. The training inputs include time, temperature, average load of previous weeks for the time of prediction, the load of prior timeslots and the loads of prior days at the time of prediction [12].

This thesis presents the results of an investigation to model the electrical energy consumption profiles associated with sanitary hot water heating in student residences using the Adaptive

Neuro-Fuzzy Inference System (ANFIS) learning method. The main application of the research is the predictive modelling of adjusted baseline energy consumption for measurement and verification purposes, which typically involves performance evaluation of energy management (EM) interventions. The model is implemented on the MATLAB platform, with the view to explore optimal configurations of differing inputs to the ANFIS method. The effects of compartmentalising the dataset into subsets representing different characteristics, thereby deriving different models representing different cyclic periods, are explored.

1.2 Project motivation

1.2.1 Overview

This section aims to elaborate on the motivations that drive the research. Two possible applications for the proposed ANFIS model are given, namely medium-term load prediction for planning and budgeting purposes and the measurement and verification (M&V) performance evaluation of energy management interventions. Thereafter follows a motivation regarding the proposed methods in which these models are to be constructed. The reasoning behind the project motivation is visualised in Figure 1-1.

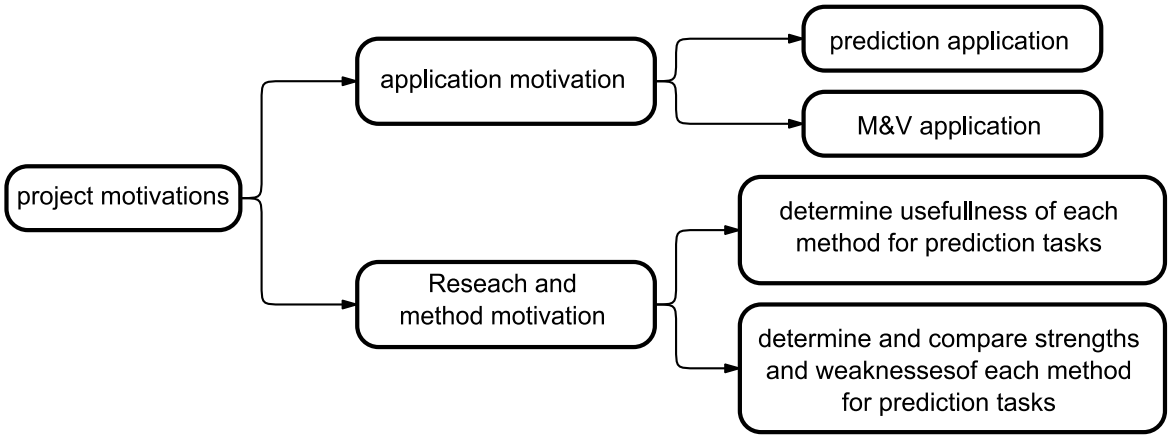

Figure 1-1: Diagram illustrating the facets of the motivations that are to drive the research.

1.2.2 Medium-term load prediction

An electrical utility should only ever produce the amount of power that is being consumed by the grid as a whole at that point in time [13]. The problem with this situation is that the demand for electricity can vary significantly. While it is necessary for certain power plants of the utility to be operational at all times to maintain the minimum possible demand on the grid,

called the base-load, the fluctuations in the demand require other plants to be dispatched to supply load following and peak load components [14]. Modern power plants that are generally expensive to build but relatively efficient and cheaper to operate, such as coal-fired, nuclear and hydroelectric plants, are typically used for baseload generation. Plants that are relatively inexpensive to build but very expensive to operate, such as simple cycle gas turbines, are used only for peak generation [14]. The plants that fall between these two categories are called intermediate load or mid-merit plants and are dispatched as needed to supply the load following component [14], [15]. Different power plants have different fixed costs and variable costs, i.e. costs that would be borne regardless of whether the plant is running and also different operational costs respectively. The utility aims to run combinations of these plants to suit the real-time demand in a way that makes economic sense. A load model of residential power use, which draws on data such as future weather predictions, can be used to plan and optimise the operational schedules of power plants [12].

In the context of the energy consumption of hot water loads, medium-term predictive modelling is useful for applications such as predicting the medium- and long-term demand growth, the savings potential of Energy Conservation Measures (ECMs), budgeting for energy costs, etc.

1.2.3 Measurement and Verification performance evaluation of energy management interventions

In general it can be said that the purpose of M&V is determining the success of energy efficiency and demand-side management (EEDSM) interventions, which aim to increase energy efficiency and thereby lower the baseline utility bill of the client [16]. In order to determine the success of these projects, it is necessary to quantify the savings impacts thereof using methodologies that yield repeatable results which is accurate to a known degree of uncertainty. The savings are represented by the difference between the electricity consumption after implementation of the intervention and the consumption that would have been the case without implementation of the EEDSM intervention [16]. The electricity consumption before implementation of the intervention is used to derive a baseline, a representation of electricity use under a set of service conditions under which the system in question was operating [16]. These service conditions can include energy governing factors such as building occupancy, environmental conditions, tariff structure, etc. that have an impact on the electricity consumption [16]. To compensate for any changes in the service

conditions, service level adjustments (SLA) are implemented, which entails adjusting the pre-implementation baseline in such a manner that it represents the post-implementation electricity consumption without the intervention [16]. It is in this baseline adjustment and calculation process where it becomes evident that a predictive load model, utilising measured values of chosen operational conditions in its training process, could prove useful in performing automated SLAs.

1.2.4 Potential of Adaptive Neuro-Fuzzy Inference Systems for water heating energy consumption modelling

Limited data is available on the daily hot water consumption profiles for the various load sectors. A fair amount of research has been conducted into the total daily hot water consumption per person for a number of consumer classes [6], [5], [7]. Water heating, however, represents a substantial part of the energy consumption of typical households [17]. It is therefore expected that daily hot water consumption profiles and the associated electricity consumption profiles will align closely with the prevailing residential electricity demand profiles. Although residential electricity demand varies, it does follow certain trends. Repetition on these trends occurs on a time of day, time of week and time of season basis. On a daily basis, the residential demand tends to start off low in the early morning, followed by a peak in the late morning, as well as a peak in the early to late evening [11]. The demand then drops significantly towards the end of the day and early hours of the following morning [11]. The residential demand profile follows a weekly trend in the sense that the weekdays exhibit a higher usage profile compared to the weekend [11]. There are seasonal trends in the demand pattern, but whether demand is higher in the winter or summer seasons depend on the heating and/or cooling requirements associated with the geographical location [11], [14]. Some of the exogenous factors that influence the varying behaviour of these trends include atmospheric conditions such as temperature, rainfall and precipitation, as well as working hours [11]. These factors influence human behaviour and therefore has an impact on how consumers use electricity to alter their surroundings. It has been shown that at least some of these exogenous factors, such as air temperature, have a nonlinear relation to the daily residential demand [9], [18].

So the problem statement with load modelling is how to organize data into patterns according to certain traits, and then predict the load behaviour using known independent variables that correlate with previous data characteristics. A collection of useful computing tools such as

soft computing and clustering, that potentially represent viable solutions to this problem, have been developed over the years. The term soft computing describes the collective or singular use of many different computing methods such as Artificial Neural Networks (ANN), Fuzzy Inference Systems (FIS) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and many more [19]. By its nature, soft computing is tolerant of imprecision and uncertainty and does not require exact input to output matching. This makes it ideal for the load model problem, where correlation between the power consumption and information on variable external factors may be vague and nonlinear [19]. Complimentary to this, clustering is defined by Jain et al, as "*the unsupervised classification of patterns (observations, data items, or feature vectors) into groups (clusters)*", and while not easily usable as a standalone classification and prediction tool, clustering processes are inherently used within certain soft computing methods to generate a starting rule base on which optimisation permutations may be performed [20].

1.3 Project description

1.3.1 Overview

In this section the project is described in terms of the research objectives, key questions regarding the problem that need to be considered, and the research tasks deemed necessary to complete the research in the given context.

1.3.2 Research objectives

There has been success in the field of residential load modelling using soft computing methods such as ANN and ANFIS [11], [18]. The objective of this project is to investigate the suitability of ANFIS for modelling the energy consumption profiles associated with water heating in student residences over differing time periods, determining which external variables will contribute to successful predictions and then comparing the accuracies of the different models using various accuracy assessment metrics. The proposed load models are to be constructed in multiple configurations of the ANFIS function, not only by varying the inputs of external variable data fed into each model, but by partitioning the load model into smaller models that deal with subsets of the given data. This is expected to give comparative insight into the ease of implementation, the computational requirements, the range of input variables required, and the strengths and weaknesses of the implemented models. It should

also provide insight into the feasibility of any real world practical implementations of these methods.

The following research objectives were identified:

- Determine the feasibility of the application of ANFIS to model and predict the hot water energy consumption profiles associated with student residences for medium term forecasting horizons.
- Perform a comparative study in which different implementations of the ANFIS function are utilised with the objective of gaining insight into the relative strengths and weaknesses of the various implementations. This objective entails the following:
  - Determine the effect of differing inputs to the load models.
  - Determine whether partitioning the load model into multiple models that each operates on a subset of the given data improves accuracies and reduces the computational complexity and runtime by eliminating certain inputs.
- Determine the feasibility of any real world practical applications for modelling techniques described, with reference to:
  - Time-series forecasting of the expected average load, where short term forecasting accuracy is expected.
  - Predictive calculation of total energy consumption within specific medium term horizons.

1.3.3 Key questions

The following key questions concerning the problem are posed:

- What is the scope of the research to be conducted, i.e. what specific soft computing methods should be researched and considered for implementation?
- What sample data should be used and why?
- Are there types of data that will not work with the proposed methods?
- Does the data need to range over a minimum timeframe and/or large enough sample area in order to be a sufficient indication of real grid performance?
- What software should be used to accomplish the following:
  - Manipulate and store the sample data.
  - Train a model by means of soft computing methods.

- Which variable factors should be used along with the power consumption profiles in order to more accurately predict trends? These may include the following:
  - Occupancy.
  - Human activity schedules.
  - Ambient temperatures.
  - Precipitation.

1.3.4 Research Tasks

The research objectives give rise to the following research tasks:

- Perform a literature review.
- Obtain suitable load test data that is representative or similar to a residential load.
- Obtain external factor data that spans over the same time as the load data.
- Choose software with which to manipulate data.
- Construct a relational database in which to store the data.
- Interface data manipulation software with the analytical software for exporting purposes.
- Sanitise, validate and combine all data that is to serve as possible input for model training into datasets.
- Write plotting functions that can adequately visualise the different components of the datasets. This will also be used to visually inspect model outputs.
- Write the program or script that will do the following
  - Sort the input datasets into training and testing datasets.
  - Train the machine learning routines with the input dataset.
  - Test the accuracies of the routines with different methods using the testing datasets.
- Set up a testing regime for each model in order to obtain the optimal parameter values for each task.
- Perform case studies with different inputs of external variables, to highlight redundant factors and test the accuracy of each attempt.

1.4 Research methodology

This research methodology, or experimental procedure, adopted for this project derives from the classical statistical model design [21]. This approach is directly applicable to the empirical machine learning procedure and is implemented to ensure model validity through rational scientific experimental design. The seven steps of this methodology are listed below, along with a short description regarding the purpose and procedure applicable for each:

- *Statement of the problem*: The problem is clearly stated and important domain specific knowledge of the research area concerning the problem is discussed and elaborated upon. The learning method or modelling technique used is irrelevant, this step focuses purely on the application domain of the modelling problem.
- *Hypothesis formulation*: Model inputs and outputs are specified, and the impact of each, which is to be established through experimentation, is hypothesised.
- *Case study data acquisition*: Data that is relevant to the domain of the problem statement of step 1 is sourced, and presented as the primary case study of this thesis. Details are given on the origin and format of the data. This data includes the relevant information to satisfy the inputs and outputs deemed necessary in step 2.
- *Data handling & pre-processing*: The data handling portion of this step consists of initial formatting and storage of the data, including some detail on the design of a relational database. Thereafter input and output datasets are constructed and pre-processed. The pre-processing steps are model specific and cater to the requirements of the ANFIS method, and include outlier removal, appropriate variable scaling, and variable encoding.
- *Experiment design*: The models that utilise the inputs specified in step 2 are constructed. The different models are designed using different permutations of the input types and rule extraction methods available in ANFIS, as well as differing combinations of optional inputs. The experimental logic behind the design choices is presented.
- *Model estimation*: The models are trained with designated training input and output sets, and then their accuracies are determined by obtaining the model outputs for testing dataset inputs, and comparing it with the actual output values using accuracy assessment methods. A comprehensive accuracy report of each model is generated in such a way that the performance of certain inputs can be compared.
- *Interpretation of the models and drawing conclusions*: The models are interpreted by observing performance changes in prediction accuracies between models, with the use of

the comprehensive reports mentioned in step 6. From this the best performing model designs are obtained. A case study on the measured real world load data is presented using these models.

The steps outlined in the methodology description are not necessarily addressed in the specific sequential order shown. This is because different components of each step may need to be addressed in different chapters. Table 1-1 indicates specific sections of the document that relate to the components of the research methodology.

Table 1-1: Guide to which sections of the document specifically relate to steps of the chosen research methodology.

Research methodology component	**Section relating to component**
Statement of the problem	1.1,1.2 and 3.1
Hypothesis formulation	3.4
Case study data acquisition	3.2 and 3.3
Data handling & pre-processing	3.5 and 4.2
Experiment design	3.7
Model estimation	5.4, 5.5, 5.6 and 5.7
Interpretation of the models and drawing conclusions	5.8, 5.9, 5.10 and 5.11

1.5 Overview of report

The remainder of the document is structured as follows:

- *Chapter 2: Literature Review:*

  A literature review is presented on relevant research pertaining to hot water consumption, the principles of machine learning, load forecasting and ANFIS modelling.

- *Chapter 3: Model Implementation:*

  In this chapter the modelling problem is explored in detail with regard to the available data, the possible model inputs, necessary pre-conditioning of inputs, the statistical properties of the training data and the design of a case study testing methodology.

- *Chapter 4: Software Implementation:*

  Software implementation entails the pre-processing of the datasets, the development of the relational database, and the implementation of various arrangements of load model prediction methods using the ANFIS functionality of MATLAB.

- *Chapter 5: Modelling Results and Performance Evaluation:*

The different arrangements of load model configurations and input types are implemented using real load and weather data, along with other information such as holiday and examination dates. The results are discussed and compared.

- *Chapter 6: Conclusions and Recommendations:*

  Final conclusions and recommendations are made and suggestions are presented for further work in this field.

Chapter 2: Literature Review

2.1 Overview

The research problem addressed in this investigation focuses on the implementation and performance evaluation of ANFIS models for modelling the energy consumption of a sanitary water heating load. This literature review focuses on material from literature that is most relevant to the study, including the following:

- Previous work in the field of sanitary hot water consumption.
- The principles of machine learning applied in the study.
- The ANFIS theoretical framework.
- ANFIS modelling in the Matlab environment.
- Prediction accuracy metrics.
- The application of ANFIS in literature for load forecasting.

2.2 Sanitary hot water heating

2.2.1 Overview

This section presents the findings of previous work in the field of analysis of sanitary hot water consumption rates of different load sectors, with an emphasis on heat pump technology. The behavioural differences of hot water consumption profiles throughout seasonal, weekly and diurnal cyclic periods are highlighted.

2.2.2 Diurnal cycles in sanitary hot water use in different load sectors

There is a difference in the pattern of daily sanitary hot water use in different load sectors. These patterns depend on the routine of the inhabitants within each sector, which are as follows:

- *Households in the residential sector:* The energy requirements of water heating installations for residential use are mainly concentrated in two peak usage periods throughout the day [5]. The first of these occurs in the morning between 06:00 and 08:00 and the second between 19:00 and 21:00. Per capita power consumption is highest in low density residential areas, lower for medium density areas and lower still for high density areas [5]. The peak of the high population density usage occurs earliest at 06:00, the

medium at 07:00 and the high at 08:00 [5]. Reflecting the difference in working hours of the respective economic classes that occupy each region. The second daily peak is usually as high or slightly higher than the morning peak [22].

- *Labour accommodation in the industrial sector:* The load associated with the heating of water for sanitation in the worker accommodations common to the industrial sector does not have the same two-peak daily cycle as with residential households. Instead the load consumes power at a constant rate throughout most of the day from 06:00 to 20:00 [6]. This is because of the daily workload is commonly divided into two or more shifts, which spreads the usage profile out over the range of shift cycles of the workers [6].
- *Hospitality industries in the commercial sector:* This is an example of how the daily hot water power use profile can be affected by something other than personal sanitation related loads, as it was found that hotels have a very large peak in power usage in the morning between 06:00 and 09:00 [6]. This peak as a result of the daily operation of the washing machines used to wash the laundry for every used room [6].

2.2.3 Weekly cycles in sanitary hot water use

Consumption of hot water for residential installations is found to be lower for Saturdays and Sundays than the days of the work week [5]. The morning and evening peaks also occur slightly later in the day. This could be because fewer people work over weekends, resulting in a more varied wake up time across a population, which in turn leads to a lower peak [5].

2.2.4 Seasonal effects on average heat pump power consumption

There are two factors that contribute to heat pumps using more energy in the winter than in the summer:

- Heat pumps work through heat exchange of air-to-water heat with a fan-coil evaporator and a water cooled condenser, therefore greater ambient temperatures characteristic of summer months result in a higher efficiency of heat exchange from the air [23], [24]. Therefore less power is used.
- The more prominent reason is that consumption of hot water for sanitary purposes is much higher in colder circumstances [6].

2.3 Theory of machine learning elements

2.3.1 Introduction and scope

ANFIS represents a modern machine learning methodology and the most relevant theoretical aspects of machine learning theory pertaining to the study are briefly introduced below. Machine learning can be defined as the process of constructing computer programs that automatically improve with experience, and makes up part of the very broad field of learning from data [21], [25]. A learning method is loosely defined as an algorithm that maps and quantifies the dependencies between the inputs and outputs of a system. Along with machine learning, other learning methods include pattern recognition and pure statistical approaches such as regression and classification [21]. While there is common ground between these methods, they are from different fields of science and differ greatly in their approach [21]. This is an important factor when considering the scope of the research.

The modelling strategy adopted for the research use the ANFIS learning procedure, which utilises the compound implementation of fuzzy inference systems as well as artificial neural networks, both of which are pure machine learning techniques [10]. Statistical theory alone does not generally provide for flexible estimation with finite samples as it favours large parametric datasets, whereas machine learning and artificial intelligence methods favour empirical approaches, with the aim of deriving complex nonlinear dependencies between inputs and outputs, while remaining relatively tolerant to imprecision [21], [25]. It is for these reasons that the scope of this literature review focuses on the pure machine learning methods that ANFIS is derived from, as well as the ANFIS method itself. That is not to say that statistics will be excluded in its entirety, as empirical methods have inherent statistical limitations, and the research methodology of the overarching practical work itself follows the general experimental procedure adopted in classical statistics, described in chapter 1.4 [21].

2.3.2 General principles

2.3.2.1 Model input and output types and terminology

The variable types, terminology, and notation used with learning model problems are the same as with classical statistics. Input and output variables have three basic types, and output variables are denoted differently depending on the type. The types, namely quantitative, qualitative, and ordered categorical types, can be described as follows [26]:

- *Quantitative:* Some measurements are bigger than others, and those that are close in value are close in nature. Quantitative inputs are denoted with X and outputs with Y. The content of both are expressed in standard set theory notation [26].
- *Qualitative:* This type implies there is a finite set of possible classes. It can therefore can also be referred to as categorical or discrete variables [26]. There does not need to be an explicit ordering in the classes. The inputs are once again denoted with an X, while the output is denoted with G. Descriptive labels, also known as linguistic variables, or numeric codes can be used interchangeably in most cases to denote the classes. An example of descriptive label use for a qualitative output is eye colour types, where the output set is shown as G = {blue, brown and green}. With the descriptive labels substituted for numeric codes it is shown as G = {1, 2, 3} [26].
- *Ordered categorical:* As with qualitative there is a finite set of possible classes, but there is an ordering between the values and no metric notion is appropriate. A simple example of categorical classes (expressed in linguistic terms) would be small, medium or large [26].

2.3.2.2 Supervised and unsupervised learning

Most learning problems can be separated into the following two categories:

- *Supervised learning:* The goal of supervised learning is to predict an outcome measure or measures based on a number of input measures [26]. During the training of this type of learning, the model requires knowledge of what the result or outputs should be, i.e. what inputs cause certain output behaviour [27]. This is the classical type of learning for predictive models, such as regression models, ANN, ANFIS, support vector machines and many more [26], [27].
- *Unsupervised learning:* The goal of unsupervised learning is to describe the associations and patterns among a set of input measures and therefore there is no output measure as with supervised learning [26]. Some examples of this approach include clustering, principle component analysis, and self-organizing maps, all of which can assist in gaining empirical knowledge of the data. In many cases these act as pre-processors for supervised learning inputs [26], [27].

2.3.2.3 Classification models and prediction models

Whether a model is defined as a classification or prediction model is primarily determined by the type of its output variable. A classification model predicts the categorical class label for a

given set of input data and the output can therefore either be qualitative or ordered categorical [25]. A prediction model predicts continuous valued functions and therefore has a quantitative output [25]. Both of these types fall under the category of supervised learning methods.

2.3.3 Inductive and deductive learning

2.3.3.1 Inductive learning

The inductive learning hypothesis reads as follows: "*Any hypothesis found to approximate the target function well over a sufficiently large set of training examples will also approximate the target function well over other unobserved samples*" [25]. Implementing this hypothesis within a learning technique means training a model with all of the available data. If the model approximates the output function well over the observed data, it is assumed that the model will also perform well for unseen instances of the data. Figure 2-1 illustrates the selection of training and validation data from the set for the case of inductive learning.

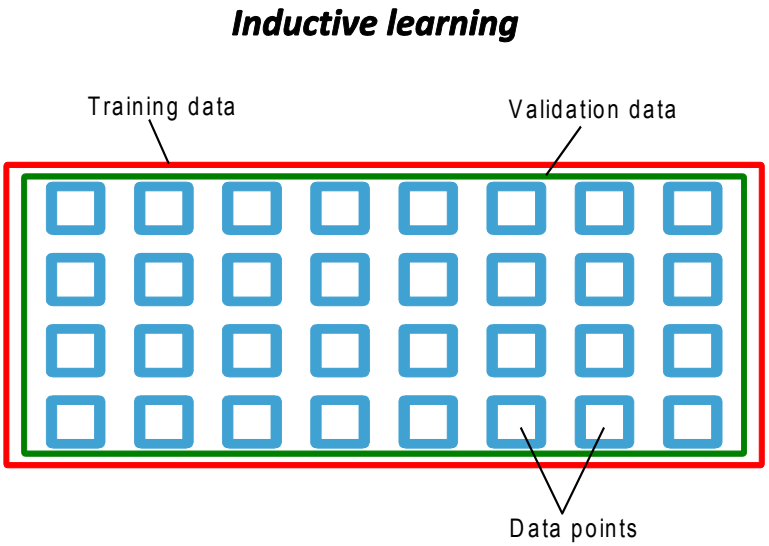

Figure 2-1: Partitioning of training and validation data in the case of inductive learning [21]**.**

There are factors that limit the reliability of this learning technique, depending on the application. The strict dependence on a very large set of training examples excludes many real world applications [28]. As is the case for many learning methods, the observed accuracy of the learned hypothesis over the training examples will typically provide an optimistically biased estimate when compared with the performance of unseen examples [25]. This bias in the estimate, combined with a rich hypothesis space, can result in over-fitting of the model to the training examples.

2.3.3.2 Deductive learning

Deductive learning is the process of partitioning the total set of available data into independent training and validation sets, as illustrated in Figure 2-2. A model is trained using the training data, and then tested using the validation data. The accuracy of the learned hypothesis over the validation examples offers a better estimate of performance over unseen examples, thereby limiting the bias in the estimate inherent in the inductive approach [25]. This does come at the expense of a smaller set of training data, but provides a mechanism with which over-fitting can be avoided.

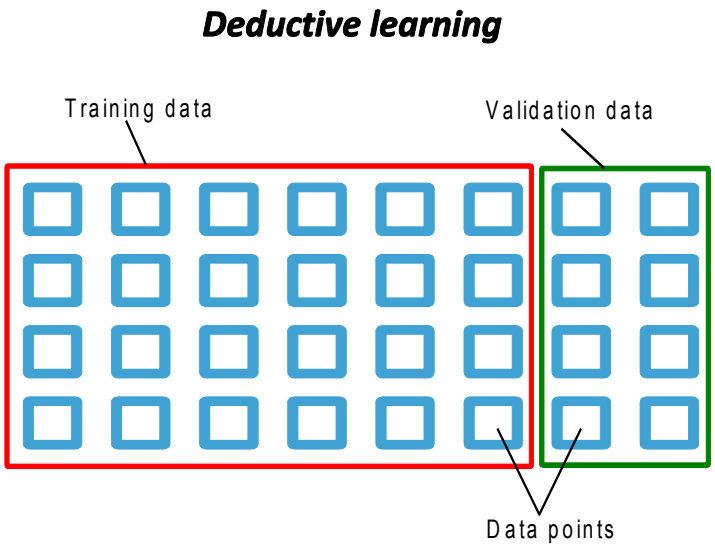

Figure 2-2: Example partitioning of training and validation data in the case of deductive learning[21]**.**

It is worth noting at this point that an estimate of accuracy derived through a deductive approach can still vary from the true accuracy, depending on the makeup of the particular set of test examples [25]. It is important to ensure that the data used is representative of all behaviours of the target concept. This variance in the estimate tends to increase with smaller training data sets [25].

2.3.4 Fuzzy systems

2.3.4.1 Introduction

The field of fuzzy systems, or fuzzy logic, stems from the introduction of the concept of fuzzy sets proposed by Zadeh [29], which was motivated by the objective to present a framework that provides a natural way of dealing with problems in which the source of imprecision is the absence of sharply defined input and output criteria for a problem. In traditional Boolean logic

a statement is either true or false, while fuzzy logic offers more flexibility as it allows for degrees of truthfulness of a statement [21].

2.3.4.2 Fuzzy sets

A fuzzy set is formally described as a class of objects with a continuum of grades of membership [29]. This set is characterised by a univariate membership function (MF) which assigns to each object a grade of membership ranging between zero and one [21], [29]. In traditional set theory a given object either is or is not a member of a set, also known as a crisp set [21]. As an example, Figure 2-3 shows how boiling temperature can be represented as a function of water temperature using the following criteria:

- *A crisp value:* In this case, illustrated in (a), water belongs to the set of "*at boiling temperature*" only if its temperature is 100°C.
- *A crisp set:* In this case, illustrated in (b), water belongs to the above mentioned set if the temperature is between 80°C and 100°C.
- *A fuzzy set:* In this case, illustrated in (c), water has partial membership in the set, dictated by the Gaussian membership function shown.

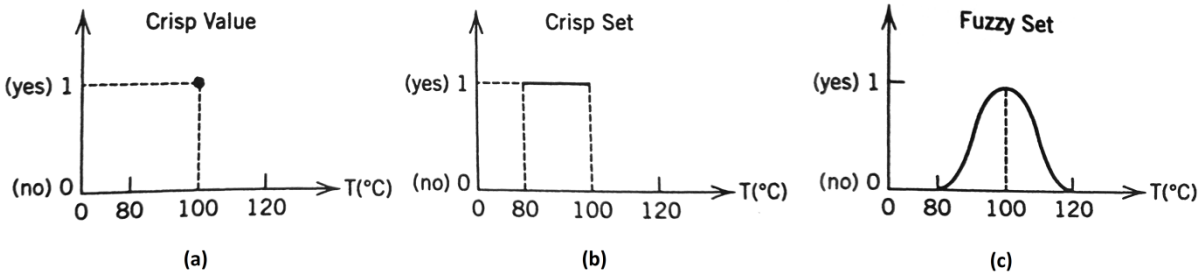

Figure 2-3: Boiling temperature as a function of water temperature membership: (a) A crisp value, (b) A crisp set and (c) A fuzzy set [21].

The mathematical properties of fuzzy sets are described in the original publication [29]. The notions of inclusion, union, intersection, compliment, relation, and convexity are extended to such sets [29]. Only the basic notational definitions of fuzzy sets are of importance for the scope of the work presented in this document. These include the following [21], [29]:

- Let X be a space of points with a generic element of X denoted by x, therefore $X = \{x\}$.
- A continuous fuzzy set A in X is specified by the fuzzy membership function $\mu_A(x)$.
- $\mu_A(x)$ associates a real number in the interval [0, 1] with each point in X.
- The value of $\mu_{v\,A}(x)$ at x represents the grade of membership of x in A.

- Therefore the value $\mu_A(x) = 0$ means that an object x is not a member of the set A, while $\mu_A(x) = 1$ indicates that x entirely belongs to A.

2.3.4.3 Fuzzy If-Then rules

Fuzzy if-then rules are expressions in the form: if A then B. Where A and B are labels of fuzzy sets that are characterised by membership functions (MFs) [10]. A rule indicates a mapping between input and output membership functions. An example that describes such mapping is: If *pressure* is *high* then *volume* is *small*. Where pressure and volume are linguistic variables and high and small are linguistic values that are characterised by MFs [30]. A Sugeno-type fuzzy if-then rule has a fuzzy premise part with a consequent part that is a nonfuzzy equation of the input parameter [31]. An example that describes such a rule in the same spirit as above is: if *velocity* is *high,* then *force = k*(velocity)²* [10].

2.3.4.4 Fuzzy inference systems

Fuzzy inference systems (FIS), also known as fuzzy rule-based systems, are a collection of fuzzy rules that specify an input-output mapping. Each of the rules represent associations between fuzzy sets in the input and output spaces [21]. The four functional blocks of which a FIS is composed, shown in Figure 2-4, can be described as follows [10], [21]:

- *Fuzzification interface*: A crisp input value is translated to a fuzzy representation and degrees of membership with the input membership functions. This step is referred to as fuzzification.
- *Knowledge base*: This functional unit is where *a priori*, or expert human knowledge, is encoded into the system. It is comprised of the functional blocks designated as the rule base and database:
- *Rule base:* Contains a number of fuzzy if-then rules.
- *Database:* Defines the membership functions of the fuzzy sets used in the fuzzy rules.
- *Decision-making unit*: Performs the inference operations on the rules.
- *Defuzzification interface*: A fuzzy output value is converted to its crisp equivalent. This step is referred to as defuzzification.

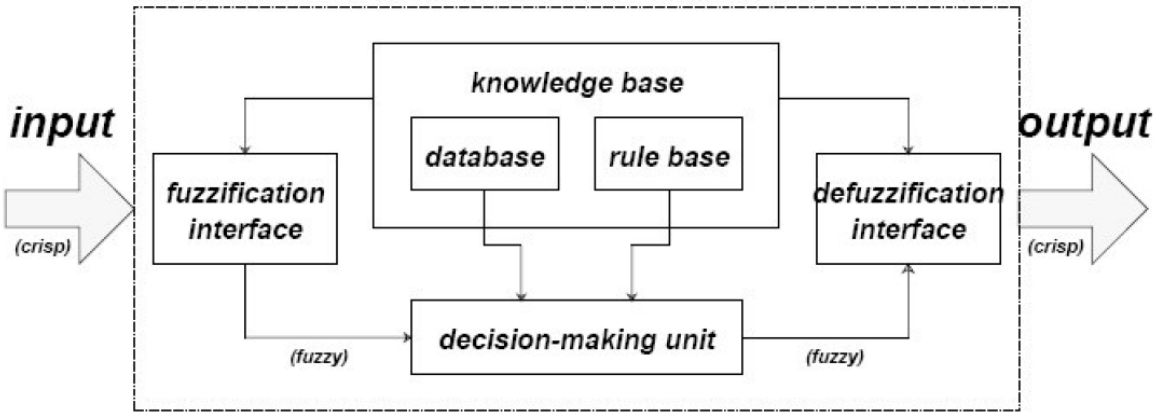

Figure 2-4: Functional topology of a fuzzy inference system [10]**.**

2.3.5 Adaptive networks

An adaptive network is described as a superset of all kinds of feed forward neural networks with supervised learning capability [10]. The network structure consists of nodes and directional links that connect functional nodes and each node performs a particular function on incoming signals [10]. The links in an adaptive network merely specify the propagation direction of a nodes output and no weights or parameters are associated with it [32]. Instead, the parameters of an adaptive network are distributed into the network nodes, each with its own local parameter set, and the union of these local parameter sets is the network overall parameter set [10], [32]. Each node in an adaptive network is either an adaptive node or a fixed node. An adaptive node, drawn as a square, has a nonempty parameter set which results in the node function being dependant on the parameters [32]. On the other hand, a fixed node is drawn as a circle and has a fixed node function as a result of an empty parameter set [32]. An example of a feed forward adaptive network is shown in Figure 2-5, with adaptive and fixed nodes and directional links that propagate exclusively from the input side to the output side [32].

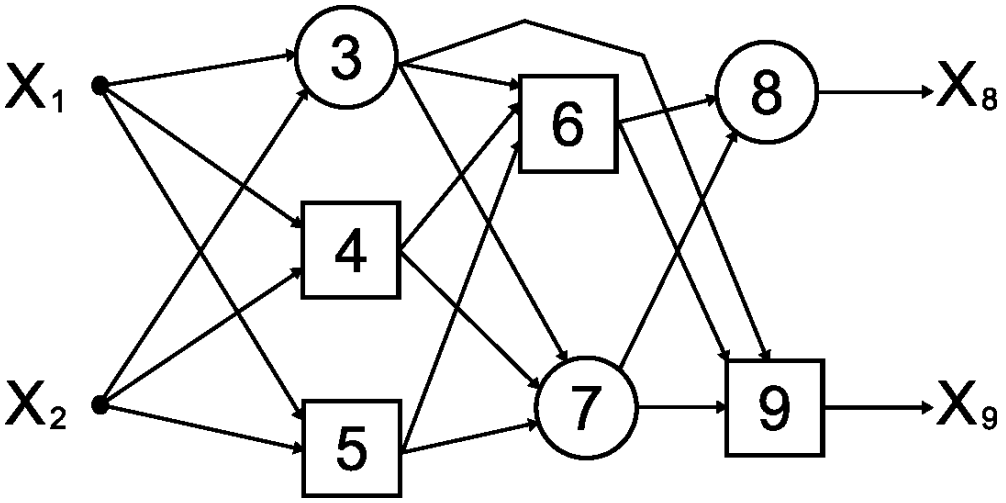

Figure 2-5: Example of a feed forward adaptive network [32].

2.4 ANFIS theoretical framework

2.4.1 ANFIS model architecture

The ANFIS architecture is typically elaborated upon in literature by means of the two-input, one-output example case. Suppose that the rule base contains the following two fuzzy rules:

- Rule 1: If x is A_1 and y is B_1, then:

$$f_1 = p_1 x + q_1 y + r_1 \tag{2-1}$$

- Rule 2: If x is A_2 and y is B_2, then:

$$f_2 = p_2 x + q_2 y + r_2 \tag{2-2}$$

The architecture of an ANFIS model with inputs x and y and one output z is shown in Figure 2-6. As with adaptive networks, square nodes are adaptive while circular nodes are fixed. The links specify the propagation direction of the previous nodes [10]. The architecture is organized into five layers and the node functions in the same layer are of the same function family. The functionality of the various layers can be summarised as follows:

- *Layer 1:* Every node i in this layer is an adaptive node with a node function described by Equations (2-3) and (2-4), where x or y can be inputs to node i (see Figure 2-6), A_i and B_i are linguistic labels for the input membership function set that they represent, and $\mu(x)$ is a specific fuzzy membership function (MF). The parameters that dictate the shape of the MF for each node in this layer are named premise parameters.

$$O_i^1 = \mu_{A_i}(x), \quad i = 1,2 \tag{2-3}$$

$$O_i^1 = \mu_{B_{i-2}}(y), \quad i = 3,4 \tag{2-4}$$

- *Layer 2:* Every node is a fixed node labelled Π, and each multiplies input signals from layer 1. The node function is described by Equation (2-5), where the output ω_i represents the firing strength of a rule.

$$O_i^2 = \omega_i = \mu_{A_i}(x) \cdot \mu_{B_i}(y), \quad i = 1,2 \tag{2-5}$$

- *Layer 3:* Every node is a fixed node labelled *N*. The i^{th} node calculates the ratio of the i^{th} rule's firing strength to the sum of all rules' firing strengths. The node function is described by Equation (2-6), where the output ϖ_i represents the normalised firing strength of a rule.

$$O_i^3 = \varpi_i = \frac{\omega_i}{\omega_1 + \omega_2}, \quad i = 1,2 \tag{2-6}$$

- *Layer 4:* Every node in this layer is an adaptive node with a node function described by Equation (2-7), where p_i, q_i, and r_i are referred to as consequent parameters.

$$O_i^4 = \varpi_i f_i = \varpi_i (p_i x + q_i y + r_i) \tag{2-7}$$

- *Layer 5:* This layer consists of a single fixed node labelled Σ that computes the overall output with the following node function:

$$O_1^5 = f_{out} = \sum_i \varpi_i f_i = \frac{\sum_i \omega_i f_i}{\sum_i \omega_i} \tag{2-8}$$

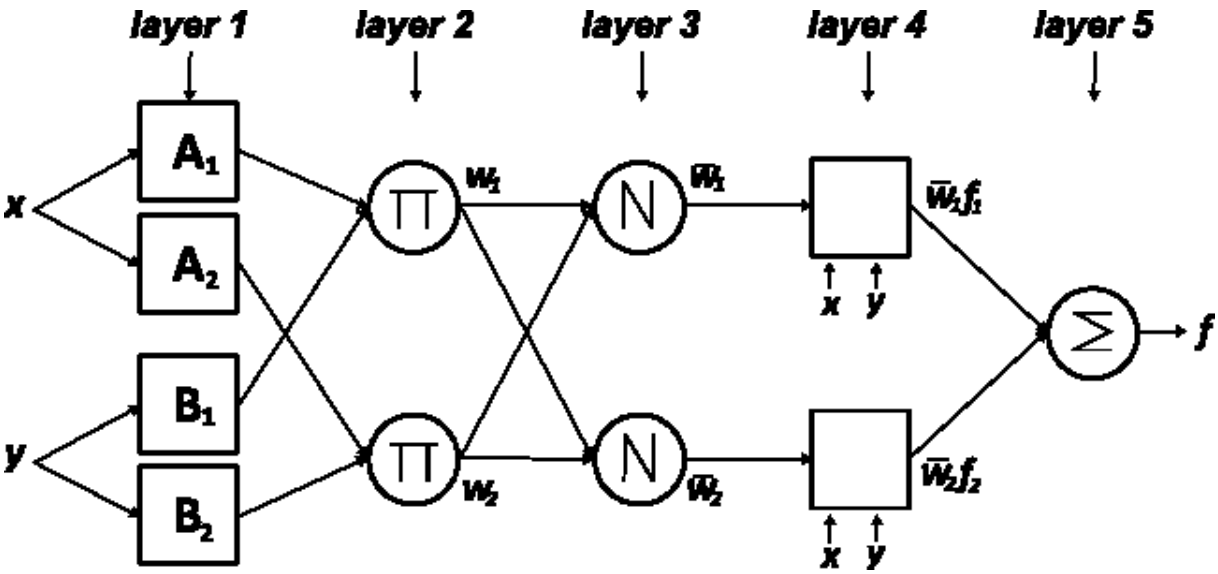

Figure 2-6: ANFIS architecture for a two input, one output inference system[30].

The number of nodes in each layer differs from this simple example depending on the number of inputs and the number of rules. Layer 1 has as many nodes as the total number of membership functions for inputs, while layers 2, 3, and 4 each have as many nodes as the number of rules used in the model [10].

2.4.2 Hybrid learning rule

The basic learning rule of adaptive networks is based on the gradient descent and chain rule, but because the gradient method can be slow and tends to become trapped in local minima, ANFIS implements a hybrid learning rule [10]. This hybrid learning rule implements both least-squares and back-propagation gradient descent methods to update the consequent and premise parameter sets during training [10], [32].

2.5 ANFIS implementation in the MATLAB environment

2.5.1 The MATLAB platform

Matlab is a platform for programming computational mathematics in a high-level language with a strong emphasis on matrix calculations [33]. Designed for engineering and scientific computing needs, it has mathematical functionalities that support linear algebra and numerical analysis as well as an extensive list of add-on toolboxes that allow for applications such as image processing, machine learning, and the fuzzy logic toolbox, etc. Furthermore Matlab has a large array of graphical tools for data visualization as well as the functionality with which to create very specific customised plotting functions.

2.5.2 Fuzzy Logic Toolbox

Matlab's Fuzzy Logic Toolbox is an add-on package that introduces functions and capabilities that allow the user to analyse and model systems based on fuzzy logic. Functions in this toolbox can be implemented using the command-line or an application. The capabilities that the toolbox adds can be summarised in three sections [33]:

- *Data clustering:* Identify natural groupings from a data set to a concise representation of the data. The toolbox provides functionality to identify clusters within input or output training data using fuzzy c-means clustering, or subtractive clustering.
- *Fuzzy inference system modelling:* The process of formulating input/output mappings using fuzzy logic to produce a FIS. The toolbox provides functionality to create both Mamdani and Sugeno FIS structures, edit and view rules, create and edit various membership function types, and generate FIS structures from data.
- *Adaptive neuro-fuzzy modelling:* An adaptive neuro-fuzzy inference system (ANFIS) is created by subjecting the membership functions of a Sugeno type FIS to the hybrid learning procedures detailed in Section 2.4.2. The toolbox provides functionality to train and test an ANFIS with separate datasets, evaluate the training errors over the range of training iterations of the learning rule with checking datasets, and save and load resulting models. Additionally, if one would like to do so, it provides the capability to change the default training options with regard to the technical workings of the learning rule, such as the number of iterations of the learning rule (epochs).

2.5.3 ANFIS functionality

2.5.3.1 Generate initial FIS

An initial FIS structure is first constructed using either the *genfis1*, *genfis2 or genfis3* functions, which take training data inputs and the corresponding output as arguments. *Genfis1* generates a FIS by performing a grid partition on the data and constructing the rules accordingly [33]. It accepts additional arguments with which the user can specify the number of MFs used per input as well as the MF types. Genfis2 generates a FIS by performing subtractive clustering of the data and constructing a rule-base that covers the feature space [33]. It accepts additional arguments that alter the clustering process. Unlike *genfis1,* the number or types of MFs cannot be specified for *genfis2*, as the number is decided by the

clustering algorithm and the type is the default Gaussian. *Genfis3* is similar to *genfis2* but uses fuzzy c-means clustering [33].

2.5.3.2 Model training

The function *anfis* is used to train a model [33]. The input arguments are the training data and the initial FIS structure. It returns a FIS with parameters tuned by the hybrid training algorithm previously mentioned, as well as the training error over the training epoch range. A checking dataset can be provided as an additional input argument so that a checking error can be tracked, which is useful for testing against overfitting of the model to the data.

2.5.3.3 Model validation

The function *evalfis* is used to validate a model [33]. The input arguments are the inputs of the testing dataset, as well as the tuned FIS structure. The output of the function can then be compared with the testing dataset outputs by means of accuracy assessment methods.

2.6 Prediction accuracy metrics

2.6.1 Overview

Evaluation of the accuracies of forecast models relies on the use of certain mathematical operations designed to give an output that is telling of the difference between a set of observed data and the set of prediction data, with a single value.

2.6.2 Error

The error is defined as the deviance between an observed value and a predicted value and it is calculated as follows [34]:

$$E = A - F \tag{2-9}$$

Where E is the error, A is the actual value and F is the forecast value.

Simply put, it is the measure of deviation of an observed value and a predicted value. This metric simply gives the size of the error, but interpretation of the error requires previous knowledge of the size of A. A more intuitive measure of the same basic metric is the Percentage Error (PE), sometimes called the relative error, which gives the error as a percentage of the actual value and is shown in Equation (2-10) [34].

$$PE = \frac{A - F}{A} \tag{2-10}$$

2.6.3 RMSE

The Root Mean Squared Error (RMSE) is a commonly used metric that gives insight into the magnitude of the errors of a series of actual and forecast values [35], [36]. It can therefore be used to assess the performance of prediction models, as each error for a series of predicted values contributes to the RMSE value. The RMSE is calculated as in Equation (2-11), where N denotes the sample size [37]. It is the metric most commonly used in training algorithms of machine learning methods, and is utilised in the back propagation algorithm used in ANN and ANFIS [10][27]. A critique of the method is that outlier errors have a greater effect on the final score, as the effect of a single error increases proportionally to its square [35].

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|A_i - F_i\right|^2} \tag{2-11}$$

2.6.4 MAE

The Mean Absolute Error (MAE) is an alternative measure to RMSE of the general magnitude of a series of errors [37]. The MAE is calculated as in Equation (2-12) [35]. Because the error is not squared, MAE is less sensitive than RMSE to large outlier errors [37].

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|A_i - F_i\right| \tag{2-12}$$

2.6.5 MAPE

The Mean Absolute Percentage Error (MAPE) is similar to MAE, but gives the error as a percentage. Therefore it is easier to interpret, as no knowledge is needed of the range of observed data sizes [34]. The MAPE is calculated as in (2-13). Because of the ease of interpretation of the MAPE metric, it is widely used in research to present the accuracy of forecast models [11], [12], [38], [39].

$$MAPE = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{A_i - F_i}{A_i}\right| \tag{2-13}$$

2.7 Forecasting with ANFIS

The following notable examples of the implementation of ANFIS for load forecasting were studied in the literature:

- Short Term Load Forecasting (STLF) of the power consumption of particular press machines in an automotive factory [1]. This particular study encompassed the design of an intelligent system that automatically configures the ANFIS algorithm by means of a multi objective genetic algorithm that automatically defines the input variables, number of membership functions of fuzzy sets for each input and the type of membership function from a pre-defined set of options for each proposed industrial load application [1]. For the press machines mentioned the inputs of weekday, hour and work shift production were automatically selected, with three Gaussian MFs each. Estimation of the loads on a quarter hourly resolution yielded low RMSE values.
- STLF of South African power networks for stable Summer and Winter months with a 30 minute prediction horizon [12]. Inputs used were the number of the half-hourly timeslot of the day, temperature, average load value of the previous three weeks at the same time of the day as prediction, the load one day prior to the same time of day as the prediction and the load 30 minutes prior to the time of prediction [12]. The inputs were assigned 12,2,2,4 and 4 MFs respectively, which resulted of a MAPE in the order of 1.37%.
- Week-ahead STLF of the half-hourly residential demand realised in France for non-transitional summer months periods of 2010 [11]. 12 inputs are used for training. 5 of these represent the previous 5 load values and the remaining 7 are load values of the same time in the previous 7 days. Week ahead forecasting is achieved by predicting the value of the following half-hour, then incorporating that result to calculate the following sample [11]. High prediction accuracy was obtained while using 4 Gaussian MFs per input.

Chapter 3: Model Implementation

3.1 Overview

The main research objective of the investigation is to determine whether an ANFIS model can predict the half-hourly averaged electrical energy consumption profiles of a sanitary water heating system for a prediction horizon ranging a few months up to one year, given appropriate informative inputs. The electrical energy consumption of sanitary hot water heating installations at student residences located at a Higher Education Institution (HEI) in the Western Cape, South Africa is used as a case study. Like all supervised learning methods, ANFIS requires appropriate selection of input and output sets of data for training, testing and validation purposes. The design of the model requires an appropriate and problem specific selection of input variables and customization of model parameters. This necessitates proper understanding of the temporal characteristics of the profiles to be modelled, as well as the factors that influence its behaviour.

3.2 Heat pump load profiles

The case study load targeted in the investigation consists of fourteen heat pumps that serve multiple student accommodation hostels on the campus of the University of Stellenbosch, located in the Western Cape of South-Africa. A numbered list of the heat pumps, along with the student residences they serve, is shown in Table 3-1. The residences accommodate of the order of 4600 students in total, which represents a sizeable statistical sample.

Table 3-1: Summary of heat pumps and residences.

Heat pump index	Residences served	Heat pump index	Residences served
1	Dagbreek, Huis Marais, Huis Visser	8	Majuba
2	Eendrag, Helshoogte	9	Minerva, Nerina
3	Harmonie	10	Monica
4	Heemstede	11	Simonsberg
5	Helderberg	12	Sonop
6	Huis Ten Bosch, Lydia	13	Tienie Louw Student Centre
7	Irene	14	Wilgenhof

The electrical load profile data acquired for the target load consists of half-hourly averaged energy consumption data for a four year period, namely 2008 to 2011. The data was obtained through using a Powerwatch™ online energy metering system. The Powerwatch™ system logs the energy consumption of the individual heat pump installations and provides a web based interface with a comprehensive set of graphing, energy budgeting and comparative ranking tools with which to analyse the data. The data for the individual heat pump installations is extracted from the Powerwatch™ database using the web interface. This yields a CSV files with forward filled timestamps.

Figure 3-1 shows the load profile of the combined heat pump load for the year of 2009. The profile reveals the following characteristics:

- Energy consumption is higher in the winter compared to the summer.
- The effects of reses periods in the academic calendar are clearly visible. These periods occur in the December-January slot, April, June-July and September. The specifics of these periods are discussed further in section 3.5.3, along with pre-processing considerations of how this impacts the modelling problem.

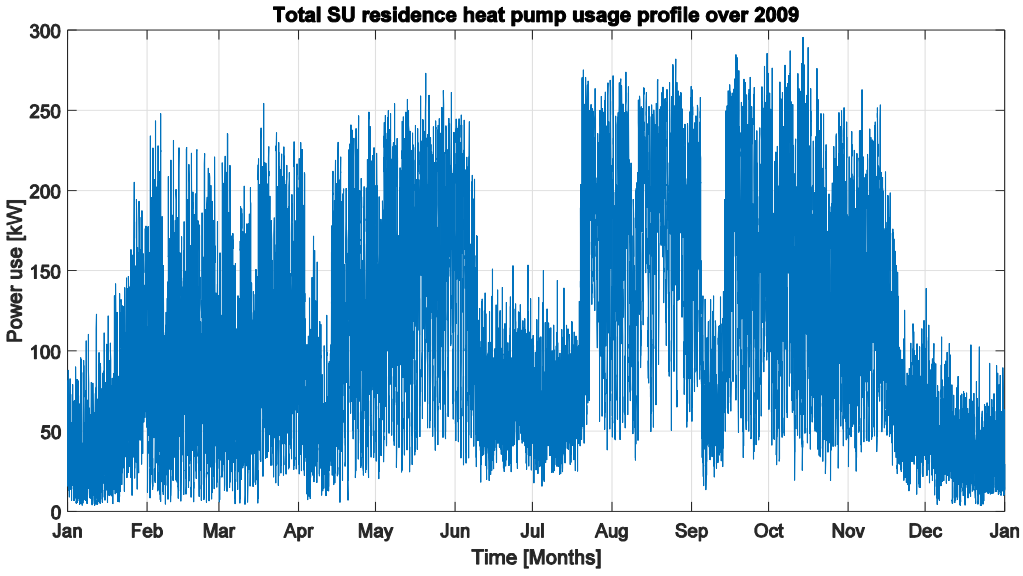

Figure 3-1: Combined heat pump load profile for 2009.

The energy consumption profiles of the targeted loads, i.e. sanitary water heating systems, reflect pronounced cyclic behaviour. These cycles are composed of superimposed seasonal, weekly and daily cycles. These cycles are typical of the cyclic behaviour observed in load profiles associated with the residential load sector.

In the context of sanitary water consumption, the annual cycle strongly reflects the seasonal climatic conditions that impact on ambient temperature. For the case study, however, time-of-year human activity such as lecturing activity, examination periods, academic reses periods, etc. also impact significantly, thereby giving rise to sub-cycles in the seasonal pattern.

The weekly cycle derives predominantly from human activity behaviour along a weekly cycle, which consists of two distinct components, namely weekdays and weekends. Mondays and Fridays typically represent transitional days, which therefore less well-defined and consistent profiles.

The daily cycle reflects the effects of night-time and day-time activities. In the morning the population awakens and may begin the day with morning rituals such as bathing, grooming, etc. This gives rise to morning in the hot water consumption profile. A second, caused by much the same reasons, occurs after the working day ends.

The above mentioned trends hold true in general but differ for certain special cases. A weekday profile will look different if it is a public holiday and may resemble that of a Saturday. Other special dates that will make the profile differ from the norm include school and university holidays, the periods surrounding important annual events such as Christmas and New Year, etc.

The load profile data used in this study spans over four years with 48 samples for each day, which yields approximately seventy thousand data points. Due to the long timelines and high resolution of the input data, it is difficult to identify cyclic behaviour and other trends therein when graphed on a single two-dimensional x-y axis. For this reason, data visualization will make use of three-dimensional graphs with colour bars and two-dimensional heat maps.

The three-dimensional plot and heat map of the half-hourly averaged power profile of the combined heat pump load for the four years are shown in Figure 3-2 and Figure 3-3 respectively. In Figure 3-2, the z-axis shows the average power while the y-axis shows the number of half hour intervals from midnight and the x-axis shows the number of days elapsed from the starting date. The yearly and daily trends in average power can be clearly observed, including the morning and evening peak and the higher power consumption for the winter months compared to the summer months. In the heat map shown in Figure 3-3, the x-axis and y-axis properties remain the same as for Figure 3-2, but the z-axis values are represented by heat map rendering. The heat map layout allows for easier comparison with other plots over the same timeframe, while lacking the visual depth of the three dimensional case.

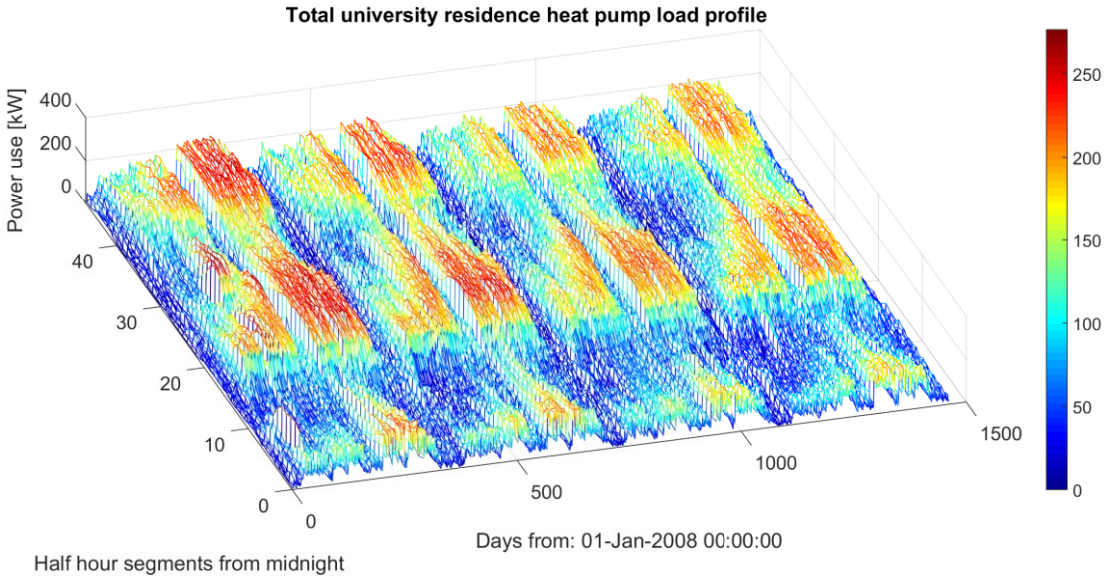

Figure 3-2: 3D plot of the combined heat pump half-hourly averaged power profile for the calendar years 2008 to 2011.

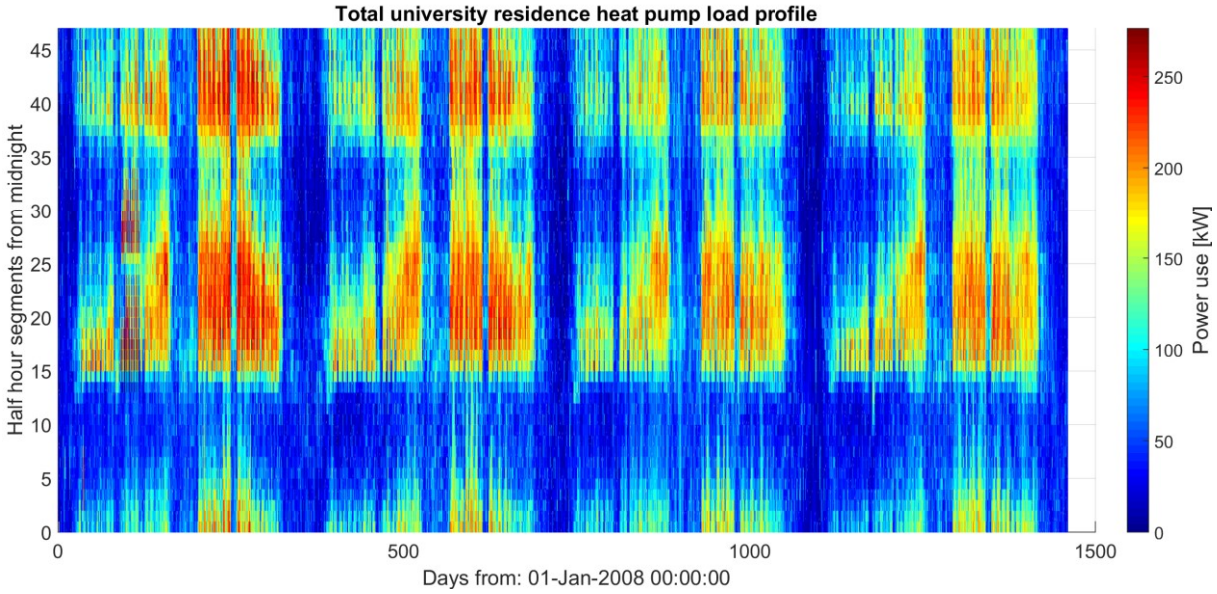

Figure 3-3: Heat map plot of the combined heat pump half-hourly averaged power profile for the calendar years 2008 to 2011.

3.3 Ambient temperature profiles

Hourly averaged temperature data for the target period, i.e. the calendar years 2008 to 2011, was obtained from the South-African Weather Service (SAWS). The data was measured at the SAWS weather station at Cape Town International Airport (CTIA), which is geographically close to Stellenbosch. The format of the original data is a structured spreadsheet, with

monthly tables structured along rows indicating the day of the month and columns indicating the hour. The tables appear in a consecutive manner, one below the other, along with descriptive information of the start date of the month for each. The average temperature is presented in backwards filled format. Figure 3-5 shows a plot of this data for 2009, where the higher temperatures associated with the summer months compared to the winter months are clearly evident.

The three-dimensional plot and heat map of the hourly averaged power profile of the ambient temperature for the four years are shown in Figure 3-5 and Figure 3-6 respectively. In Figure 3-5, the z-axis shows the average temperature while the y-axis shows the number of half hour intervals from midnight and the x-axis shows the number of days elapsed from the starting date. In the heat map shown in Figure 3-6 the x-axis and y-axis properties remain the same as for Figure 3-5, but the z-axis values are represented by heat map rendering.

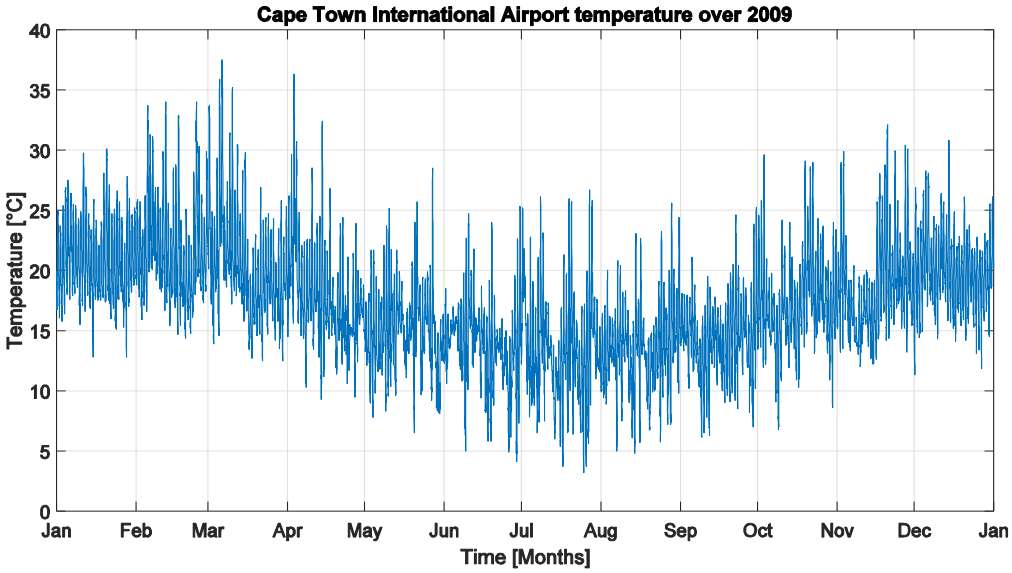

Figure 3-4: Temperature as recorded at the CTIA weather station for 2009.

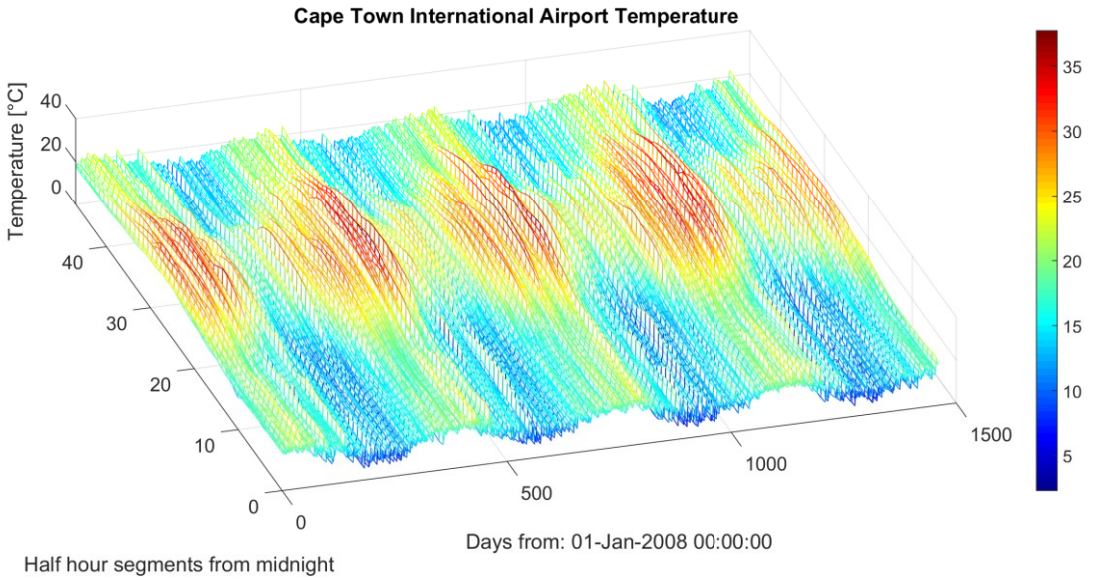

Figure 3-5: 3D plot of the temperature as measured at the CTIA weather profile for the calendar years 2008 to 2011.

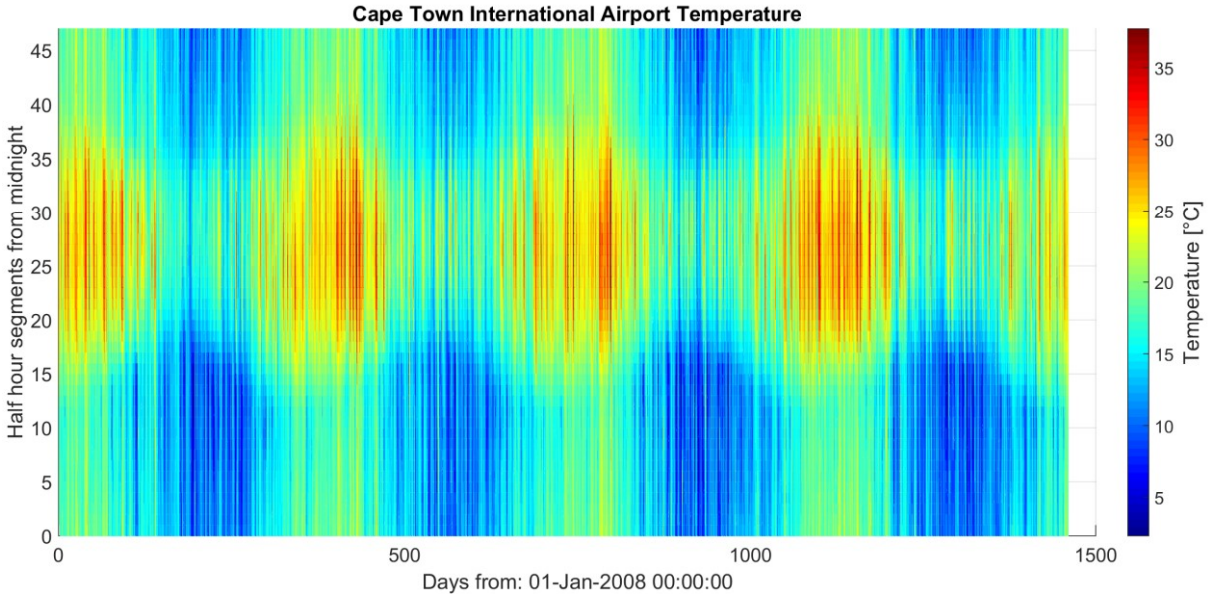

Figure 3-6: 2D heat map plot of the temperature as measured at the CTIA weather station profile for the calendar years 2008 to 2011.

3.4 Hypothesis formulation

Any deductive learning technique consists of four basic steps. The interaction of the input and output data associated with each step is described below:

- *Designation of data:* a portion of the available data, i.e. a series of descriptive inputs and the complimentary set of outputs that are associated with those inputs, is designated for training. The remaining portion is designated for validation.

- *Training:* All of the input- and output data in the training set are used to train the model.
- *Forecasting:* All of the input data in the validation set is fed into the trained load model, which then outputs the corresponding predicted load.
- *Validation:* The predicted load obtained with the specified set of validation data inputs is compared with the actual load for those inputs.

The following inputs are hypothesised to be necessary for medium term load approximation of the target data:

- *Time of day:* From Figure 3-3 it is evident that the time of the day has a very clear impact on the load magnitude. This input is given as the number of the half-hour time interval of the day. The possible values of the time of day input are 1 to 48, each indicative of the current time slot.
- *Day of week:* It has been shown that for residential sanitary hot water use that the day of the week can be an informative input, as the power usage of the days in the work week and weekend have markedly different behavioural profiles, because of the work week cycle. The input is translated to a quantitative numerical equivalent representing the seven days of the week with the numbers 1 to 7, i.e. Monday to Sunday.
- *Day of year:* The load presents seasonal variations in general power use. The day of year is a possible informative measure, as its use in training should infer periods of increased or decreased use upon specific ranges of the input. The day of year should have equal weight for each consecutive training year, as the change of the seasons should approximately occur at the same value each year. The numerical qualitative measure of this input ranges from 1 to 365 normally, but is made 1 to 366 to allow for leap years.
- *Temperature:* Through comparison of Figure 3-3 and Figure 3-6 it is clear that the seasonal change in the magnitude of the temperature is inversely proportional to the change in magnitude of the load output. This allows for experimentation as the temperature could either be used to replace the day of year input or serve as an additional input that may improve the prediction capabilities of the load model. Alternatively, it could be left out completely. Like the load, the temperature is a quantitative value with a range of 0 to the maximum value in the set.

3.5 Data conditioning

3.5.1 Equal dimensionality formatting

The format of the raw load data is half hourly with forward filled timestamps, while the temperature data is hourly with backwards filled timestamps. For the sake of equal dimensionality of the training and testing dataset pairs, the format of the load data is taken as the default. The temperature data is expanded to half hourly resolution through interpolation by means of a four point moving regression operation and the timestamp values adjusted accordingly to present the data in forward filled format. All of the graphical representations of the temperature data throughout this document (including those presented before this point) are in this interpolated format.

3.5.2 Appropriate variable scaling

Variable scaling is necessary for some learning problems in cases where different input variables have different natural scales or units of measurement and therefore can have a larger or smaller weight simply because of this. In this case the temperature data is in degrees Celsius and the load data in average kW per half hour. The input sets of each are normalized to a range between 0 and 1 by dividing each element in the set with the largest occurring measurement in that set.

3.5.3 Outlier removal

3.5.3.1 Introduction

Outlier removal entails the removal of data points within the training data that are not representative of normal behaviour for the specific modelling problem. Academic reses periods, public holidays and examination periods are examples of period that impact on the sanitary hot water consumption profile.

3.5.3.2 Deviations caused by public holidays

Public holidays that fall on weekdays result in a daily profile that more closely resembles that of a Saturday than a weekday, thereby rendering the data points of that day inappropriate for weekday training. Public holidays that fall on a Saturday go unobserved and holidays that fall on a Sunday are observed on the following Monday. Table 3-2 shows South-African public holiday dates, with the underlined dates indicating unobserved holidays. All 48 daily data

points are removed from the load and temperature datasets for each of the observed public holiday dates.

Table 3-2: Documented South-African public holiday dates throughout the load data range, underlined dates indicate unobserved public holidays.

Public holiday name	2008	2009	2010	2011
Human rights day	2-May	21-Mar	22-Mar	21-Mar
Good Friday	21-Mar	10-Apr	2-Apr	22-Apr
Family day	24-Mar	13-Apr	5-Apr	25-Apr
Freedom day	28-Apr	27-Apr	27-Apr	27-Apr
Workers day	1-May	1-May	1-May	2-May
Election day	NA	NA	NA	18-May
Youth day	16-Jun	16-Jun	16-Jun	16-Jun
National women's day	9-Aug	10-Aug	9-Aug	9-Aug
Heritage day	24-Sep	24-Sep	24-Sep	24-Sep

3.5.3.3 Deviations caused by university reses and examination periods

Most student residences are vacated during academic reses periods, resulting in periods of almost no activity in the load profile. The impact of examination periods is, however, are not that clear. When formal lectures end and examinations begin the usual routine of the student population changes. Second examination periods differ from the main examination periods because the residence occupancy is much lower. Table 3-3 shows the start and end dates of formal classes, first examination periods and second examination periods applicable for the target period. Figure 3-7 shows a heat map of the half-hourly average power profile that remains when the academic reses periods are removed. From this it can be seen that the periods before the June and December holidays exhibit very low usage as a result of the low occupancy during the second exam opportunity. Therefore the full 48 daily dataset points are removed for each day of the reses periods, as well as for the second examination periods. Figure 3-8 shows a heat map of the half-hourly average power profile that remains when the academic reses periods, public holidays and second examination periods are removed.

Table 3-3: Documented start and end dates for classes, first exam periods, and second exam periods of Stellenbosch University over the load data range.

Year	2008	2009	2010	2011
First quarter classes start	4-Feb	2-Feb	25-Jan	31-Jan
First quarter classes end	20-Mar	3-Apr	12-Mar	18-Mar
Second quarter classes start	31-Mar	14-Apr	23-Mar	28-Mar

Second quarter classes end	16-May	15-May	7-May	13-May
June exam first opportunity start	20-May	19-May	11-May	17-May
June exam first opportunity end	9-Jun	8-Jun	31-May	6-Jun
June exam second opportunity start	10-Jun	9-Jun	1-Jun	7-Jun
June exam second opportunity end	27-Jun	26-Jun	18-Jun	24-Jun
Third quarter classes start	21-Jul	20-Jul	19-Jul	18-Jul
Third quarter classes end	5-Sep	4-Sep	3-Sep	2-Sep
Fourth quarter classes start	15-Sep	14-Sep	13-Sep	12-Sep
Fourth quarter classes end	24-Oct	23-Oct	22-Oct	21-Oct
November exam first opportunity start	28-Oct	27-Oct	26-Oct	25-Oct
November exam first opportunity end	19-Nov	18-Nov	17-Nov	16-Nov
November exam second opportunity start	20-Nov	19-Nov	18-Nov	17-Nov
November exam second opportunity end	12-Dec	5-Dec	4-Dec	3-Dec

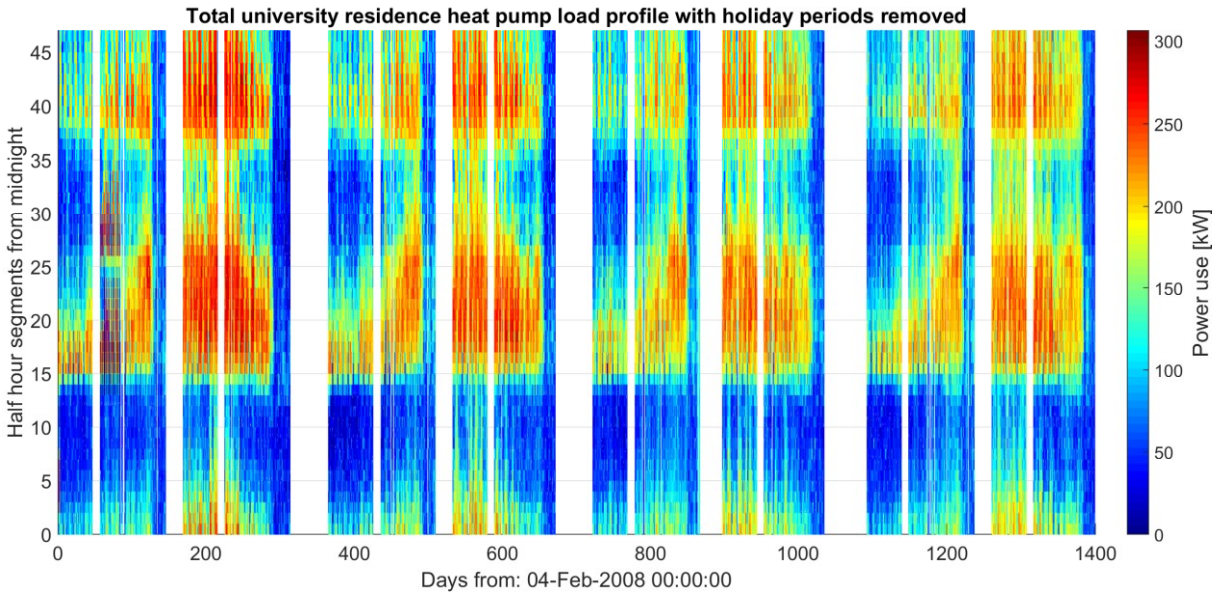

Figure 3-7: Heat map of the combined heat pump load profile the academic reses periods removed.

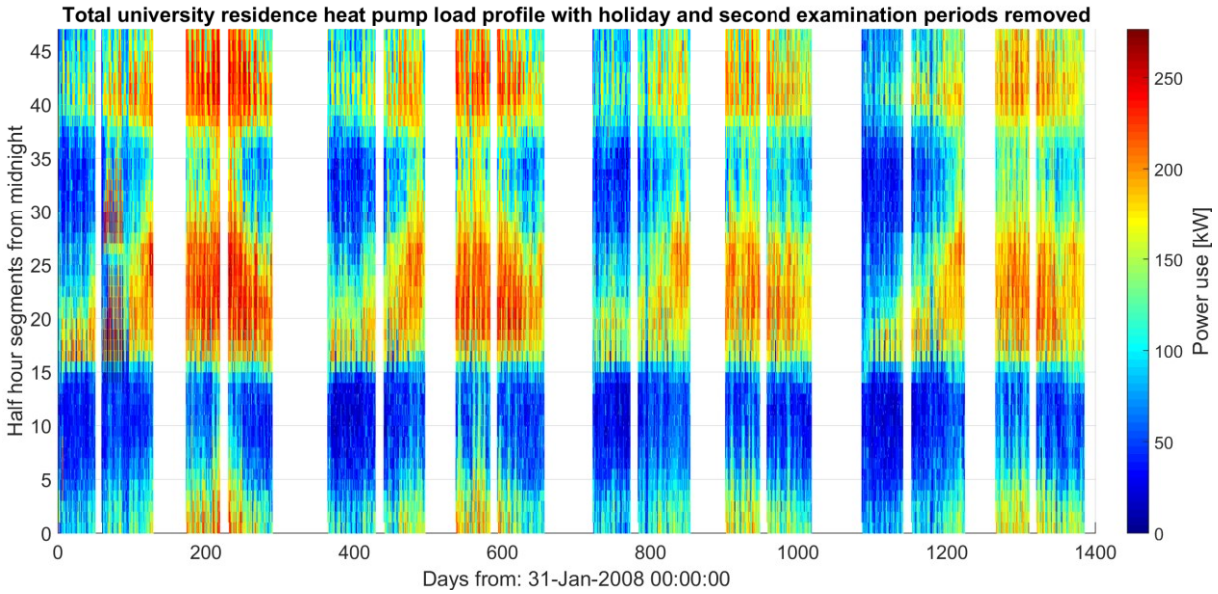

Figure 3-8: Heat map of the combined heat pump load profile with academic reses periods, public holidays and second examination periods removed.

3.5.3.4 Impacts of load shedding

There is a period of abnormal load behaviour during the second semester of 2008 where the morning peak of the total power use profile of every other day appears a few hours late, with periods of no use where the peak should be found. This can be observed in the heat map of the load data for 2008 in Figure 3-9. This is an outlier case due to load shedding events and therefore the 48 data points for each day are removed from the draining dataset for each day in the period from 1 May to 9 June. It is interesting to note that shedding this load simply shifted it to later in the day.

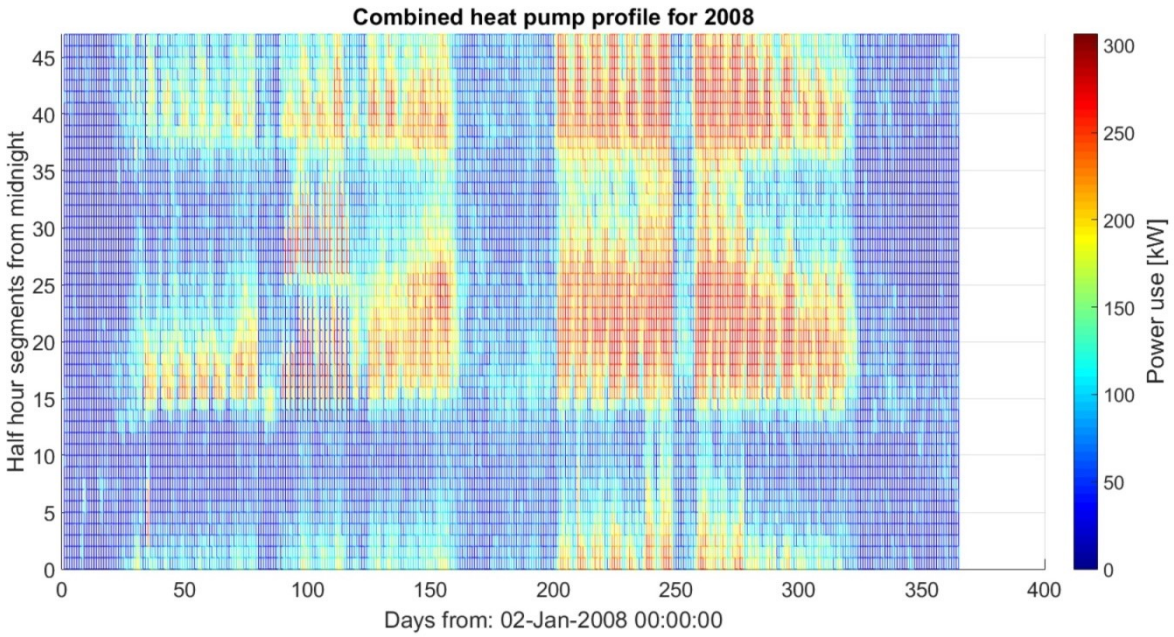

Figure 3-9: Heat map of the combined heat pump average power profile for 2008.

3.5.3.5 Inconsistencies as a result of transitional periods

The official start and end dates of classes and examinations does not necessarily imply that the bulk of students have moved in or out of residences. The first few days after the official start dates of some semesters' shows unstable and decreased usage, which could be because of students moving in after the official start date. The last few days before the end dates of examination periods show similar trends, since the final examination for many students do not necessarily fall on the final day of the examination period and therefore many students leave earlier. In cases where this transitional behaviour was found to be too far from the norm, the days where removed from the training data. Table 3-4 shows the dates removed, as well as an indication of the transitional period that the day belongs to.

Table 3-4: Dates removed due to abnormal behaviour in transitional periods

Transitional period	**Dates removed**
Start of first quarter	04-Feb-2008 05-Feb-2008
End of first quarter	20-Mar-2008 03-Apr-2009
Start of second quarter	None
End of second quarter	08-Jun-2009
Start of third quarter	21-Jul-2009 18-Jul-2011

End of third quarter	05-Sep-2008 04-Sep-2009
Start of fourth quarter	None
End of fourth quarter	18-Nov-2008 19-Nov-2008 16-Nov-2009 17-Nov-2009

3.5.4 Final training data set

Figure 3-10 and Figure 3-11 show the averaged power and averaged temperature profiles respectively after data condition.

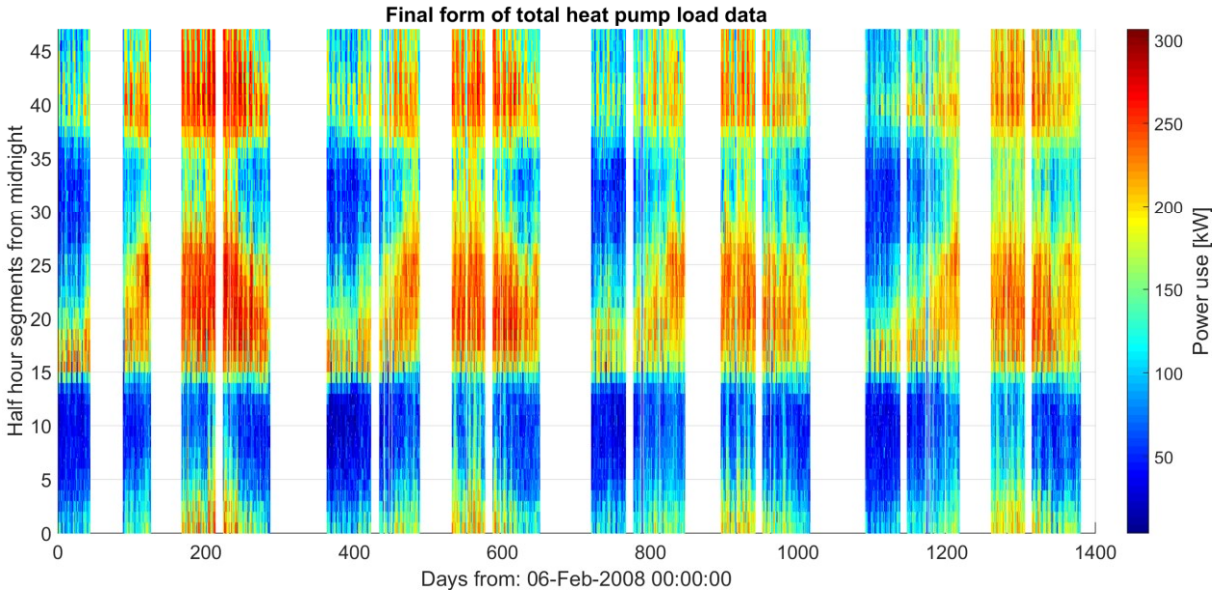

Figure 3-10: Heat map of the half-hourly average power profile of the combined heat pump load after data conditioning.

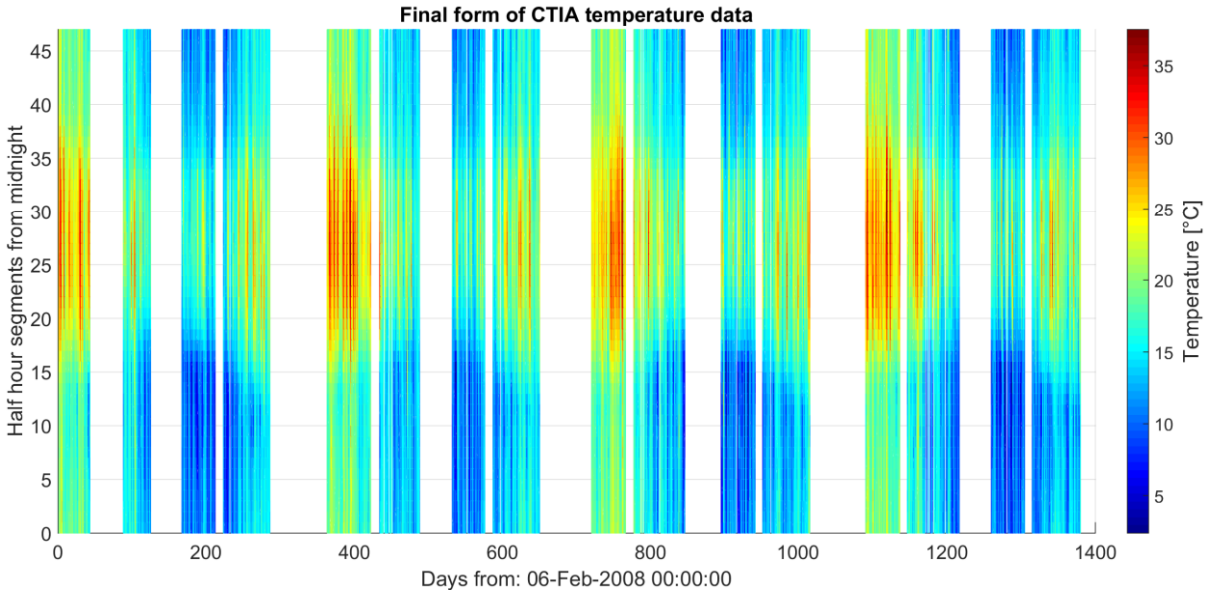

Figure 3-11: Heat map of the half-hourly averaged temperature profile after data conditioning.

3.6 Statistical analysis of training data

3.6.1 Overview

In order to evaluate the prediction accuracy of a ANFIS model, the predicted profile is compared with the observed profile over the same time period. Data from different time periods are used for training and testing. If the testing data exhibits different behaviour from that of the training data, it could have an effect on model performance. Interpretation of the accuracy results of the various models must take cognizance of any such behavioural differences. To this end, the power and temperature profiles are analysed in order to gain insight into the statistical properties that apply.

3.6.2 Heat pump load data

In order to gain insight into the behaviour of the combined heat pump load data, statistical analysis is performed on the half-hourly energy consumption profile as well as the daily energy consumption profile. The half-hourly energy consumption profile is derived from the half-hourly averaged power profile, which enables calculation of daily energy consumption data. Table 3-5 shows the yearly mean and standard deviation statistics of the half-hourly and daily enrgy consumption profiles. The mean values decreases by more than 4% per year from 2008 to 2010 and increases again by a 1.5% from 2010 to 2011. The inflated mean of 2008 might be partially ascribed to the removal of a large section of data from the second semester due to load shedding experienced during this time, which renders the data inappropriate for

training purposes. The data removed represents months with lower energy consumption compared to the winter months. The consistent decrease from 2009 to 2010, however, suggests that the energy requirements of the water heating system reduced from 2008 to 2010. This could be because of the installation of water saving shower heads in the various student residences during this time.

Table 3-5: Mean and standard deviation statistics for the half-hourly and daily energy consumption for the calendar years from 2008 to 2011.

Year	Mean half-hourly energy consumption [kWh]	STD of half-hourly energy consumption [kWh]	Mean of daily energy [MWh]	STD of daily energy [MWh]
2008	73.55	35.71	3.53	0.95
2009	70.31	33.11	3.38	0.84
2010	67.32	29.77	3.23	0.73
2011	68.29	30.15	3.27	0.83

Figure 3-12 and Figure 3-13 show histograms of the half-hourly and daily energy consumption for the calendar years from 2008 to 2011. The frequency axis values are shown as a percentage of the total. The distributions for the half-hourly energy consumption is broad and flat topped for all four years. This broad distribution is expected based on the relatively high standard deviation of half-hourly energy consumption shown in Table 3-5,. There is a yearly reduction in high consumption values from 2008 to 2010, resulting in the abovementioned lowering of the mean energy consumption and accounting for the lowering of the standard deviations. Similarly, the number of high consumption days also decreases over that time period, as is shown by Figure 3-13.

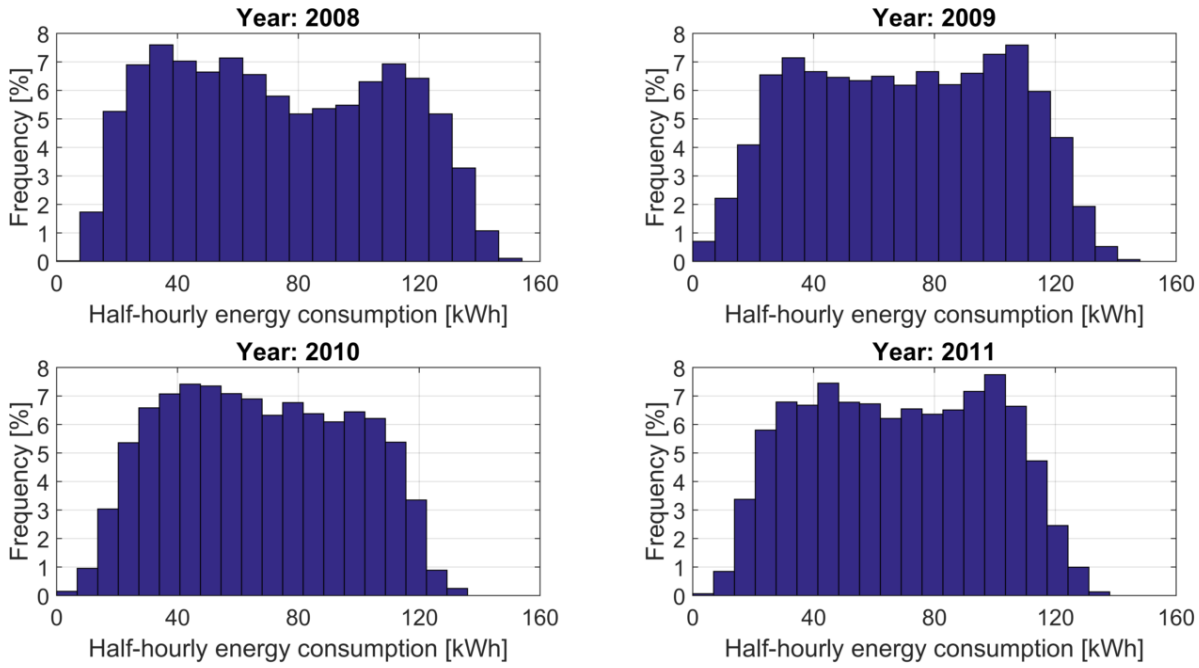

Figure 3-12: Histograms of half-hourly energy consumption for the calendar years from 2008 to 2011.

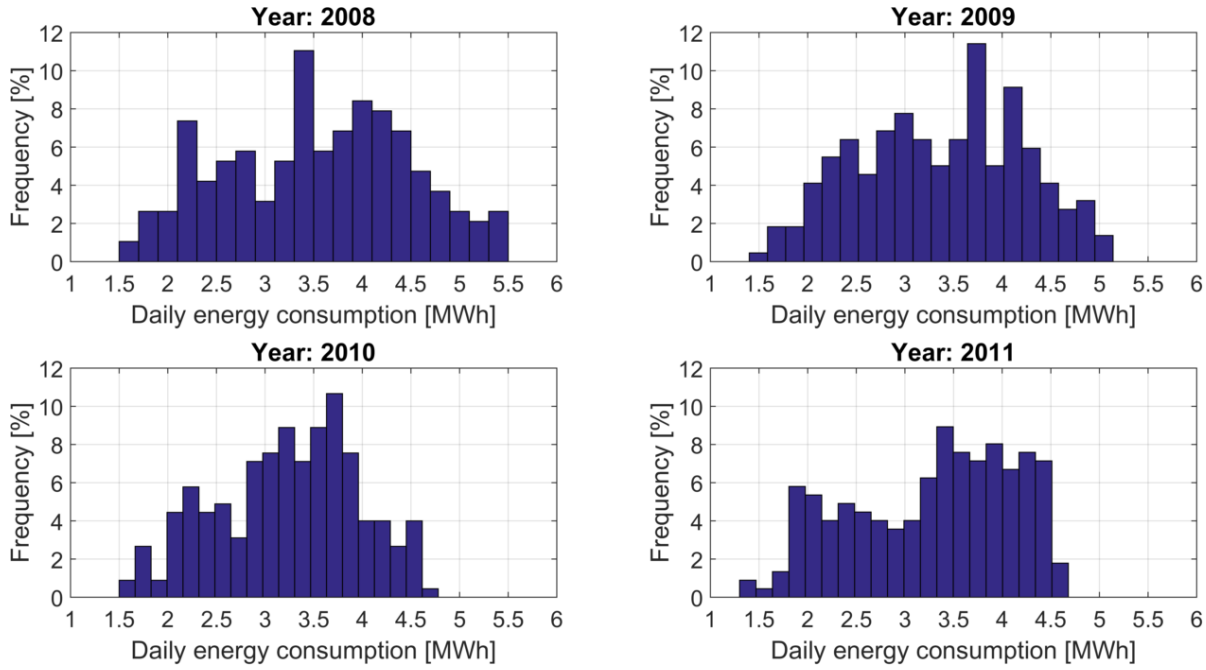

Figure 3-13: Histograms of daily energy consumption for the calendar years from 2008 to 2011.

3.6.3 Temperature data

Variations in ambient temperature over the four years can be investigated by observing the mean and standard deviation (STD) of the temperature profile. Table 3-6 shows these values for the target dataset. The results show no substantial change in either mean or standard deviation. Figure 3-14 shows histograms for each year of temperature data. The amount of

data in each year is not consistent, because of differences in data conditioning from one year to the next. The frequency axis values on the histograms are therefor given as percentage values. The results show that similar probability distributions apply for the different years.

Table 3-6: Mean temperature and standard deviation of temperatures for the calendar years from 2008 to 2011.

Year	Mean temperature [°C]	Standard deviation [°C]
2008	16.22	4.73
2009	17.02	4.72
2010	16.87	4.83
2011	16.55	5.18

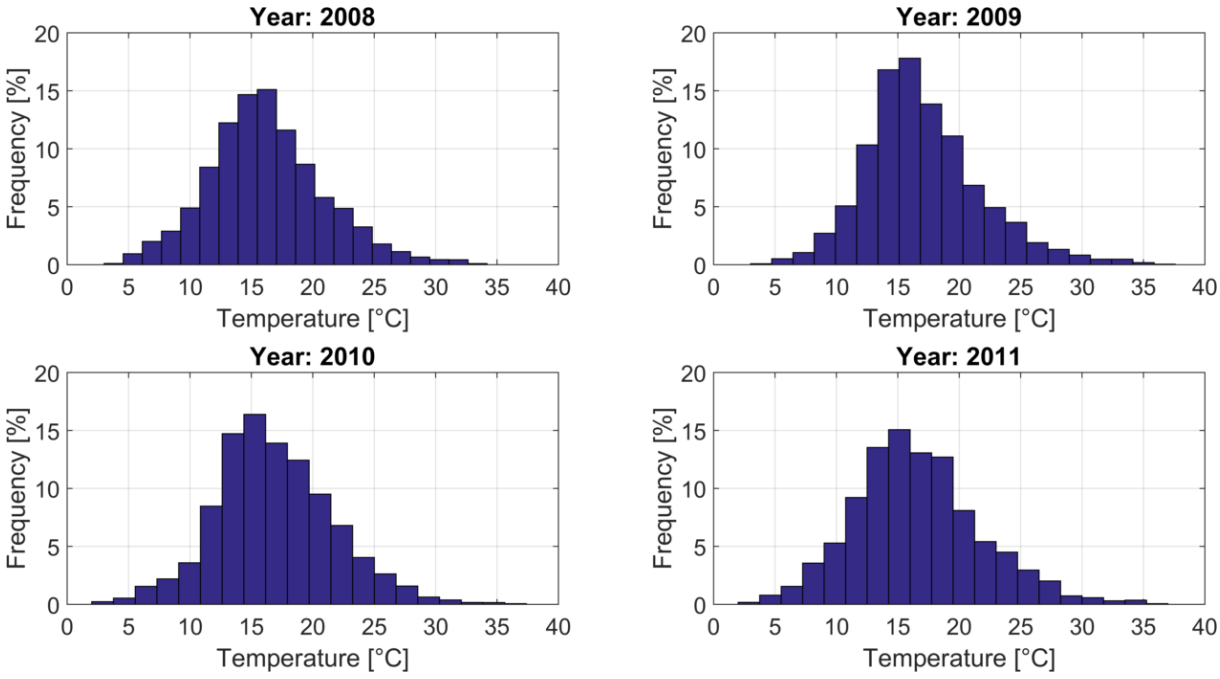

Figure 3-14: Histograms of the average ambient temperature for the calendar years from 2008 to 2011.

3.7 Experiment design

3.7.1 Overview

The ANFIS method allows for customization of the input membership functions as well as multiple means to generate an initial FIS structure. There are therefore a number of ways to configure different models for the same training data. Furthermore, non-essential inputs can be excluded from certain models so that the resulting testing outputs can give insight into the effect of those inputs on model accuracies. This section includes information on the design of

the different load models in the case study, the validation methodology and considerations of design choices regarding ANFIS training options.

3.7.2 Experimental model configurations

The base model is established and utilises all inputs specified in Table 3-7. However there is flexibility with regards to the combinations of inputs which can be used, since both the day of year input and the temperature input should be able to supply the necessary information to infer the time of the year to an ANFIS model. Table 3-7 also shows options for configurations of different input permutations. The experimental purpose of each option is as follows:

- *Option A:* Serves as the benchmark model. All inputs are used.
- *Option B:* Temperature input is excluded. The aim is to determine the effectiveness of temperature as an informative input on external factor data that has an influence on the load.
- *Option C:* Utilises all inputs but the day of year. The aim is to test the hypothesis that the temperature input can supply sufficient information to enable realistic predictions with regard to seasonal behaviour.
- *Option D:* The same inputs are utilised as for option A, only now the temperature data input is delayed by 30 minutes, so that the input to the model is the temperature during the prior timeslot of the day. The purpose of this experiment is to determine if consideration of the possible transient effects of temperature on the sanitary water heating installation.

Table 3-7: Options for alternate configurations of inputs to load model.

Training input	**Alternate input configuration**			
	Option A	**Option B**	**Option C**	**Option D**
Day of week	✓	✓	✓	✓
Time of day	✓	✓	✓	✓
Day of year	✓	✓		✓
Temperature	✓		✓	✓ (Delayed)

A further experimental consideration is the possibility of compartmentalising the model into sub-models, each targeting a specific sub-section of the data, thereby reducing the number of inputs to the load model and by extension, the mathematical and computational complexity thereof. These sub-model configurations can be implemented for any of the abovementioned set of input options. The sub-sections to be targeted, as well as the purpose of experiment are as follows:

- *Day of week:* The day of week input has 7 possible states, but it is shown in section 2.2 that the characteristics of the weekdays are similar, while Saturdays and Sundays differ. Therefore an experimental setup of 3 sub-models is proposed, with a distinct ANFIS trained to handle data from each weekday category, namely weekdays, Saturdays and Sundays. The day of week input of each data point is used to index it to the appropriate model.
- *Time of day:* The time of day input has 48 possible states, each indicative of a specific half hour time period during the day. An experimental setup of 48 sub-models is proposed, with a distinct ANFIS trained to handle data from each timeslot. The time of day input of each data point is used to index it to the appropriate model.
- *Time of day within time of week:* This experiment combines the sentiments of the first two mentioned, by constructing a distinct ANFIS for each time period of the day within each weekday category. The models are not trained with either the day of week or the time of day, as those inputs are used to index the data to the correct sub-models. Take note that input configuration options B and C will not work for this configuration, as that would result in ANFIS models with only one input, and the minimum is two.

3.7.3 Partitioning of training data

As ANFIS is a deductive learning procedure, the available input data needs to be partitioned into a portion that is designated to train the model with and a portion that will be used to validate the model. In order to achieve the medium term prediction horizon of one year, the validation data needs to span at least one year. A high training/testing data ratio is preferable. There are four years of data and at least one of them needs to be used for validation, therefore the ratio is chosen as 75%/25%, i.e. training with 3 years of data and performing validation with the remaining year of data. The implications of this possibility can be addressed by utilising cross validation.

3.7.4 K-fold cross validation

3.7.4.1 Overview

It is shown in 3.6.2 that the behaviour of the load changes slightly from one year to the next. This is not ideal for training, as the validation year may exhibit different characteristics compared to the years used to train the model. The complexity of the learning problem is further increased by the fact that the academic year, examinations and recesses do not start or

finish on the same dates for every year. The performance of a load model could be different when training with the first three years and validating with the fourth than it would have been if the first year was used for validation and the rest for training. For this reason the K-fold cross validation method is used.

The K-fold cross validation process can be summarized as follows [21]:

- The data is partitioned into K parts or folds, of which K-1 folds are selected for training, leaving the remaining fold for validation [21].
- This is repeated for K iterations, each time with a different group selected for validation purposes [21]. Since the largest of the repetitive trends in the load model problem occurs on a yearly basis, the data is separated into a fold for each of the four years of the available data.

The allocation of training and testing sets for each cross validation training iteration is shown in Table 3-8. Each case model is trained and tested K times in this manner. In this case that translates to four times more information on the performance of a load model and a difference in the performance of a model over separate folds could give valuable insights during problem solving [40].

Table 3-8: K-fold cross validation training and testing data allocation for each fold.

Iteration	Training folds	Validation fold
1	2009, 2010, 2011	2008
2	2008, 2010, 2011	2009
3	2008, 2009, 2011	2010
4	2008, 2009, 2010	2011

3.7.5 Cross validation results

3.7.5.1 Overview

The load data designated to each cross validation fold is shown in this section. This is done for ease of comparison of the load output over each fold, and to give a reference with which to compare the outputs of experimental load models presented later in this document.

3.7.5.2 Fold 1 load data for 2008

The heat map and 3D renderings of the combined heat pump power profile for 2008 is shown in Figure 3-15 and Figure 3-16 respectively.

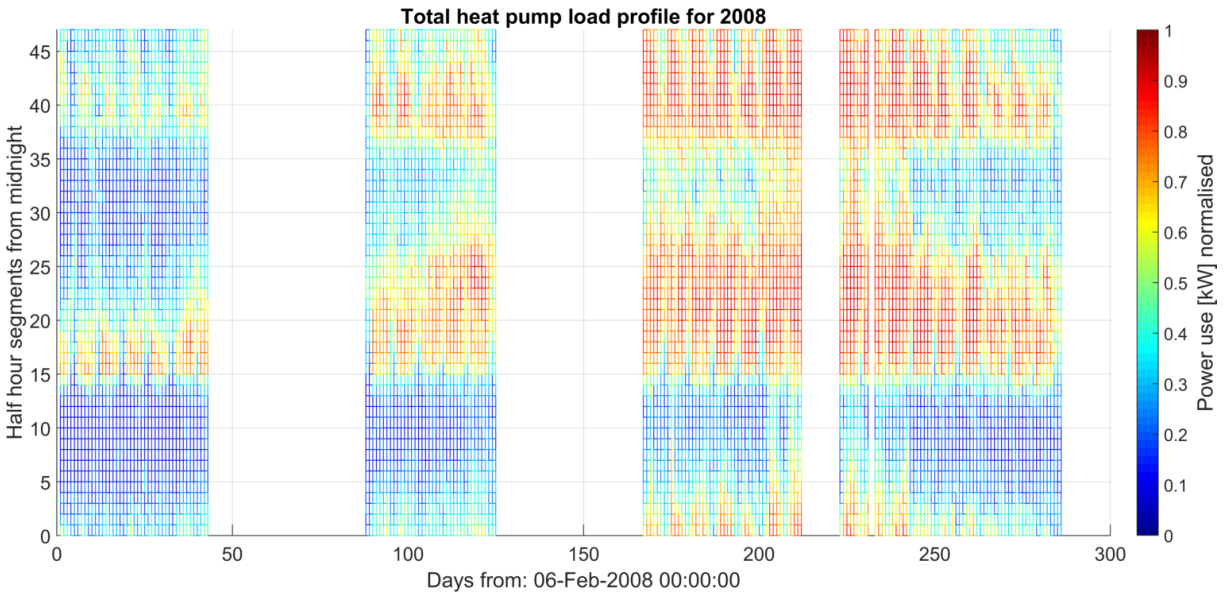

Figure 3-15: Heat map of the normalised combined heat pump power profile for 2008.

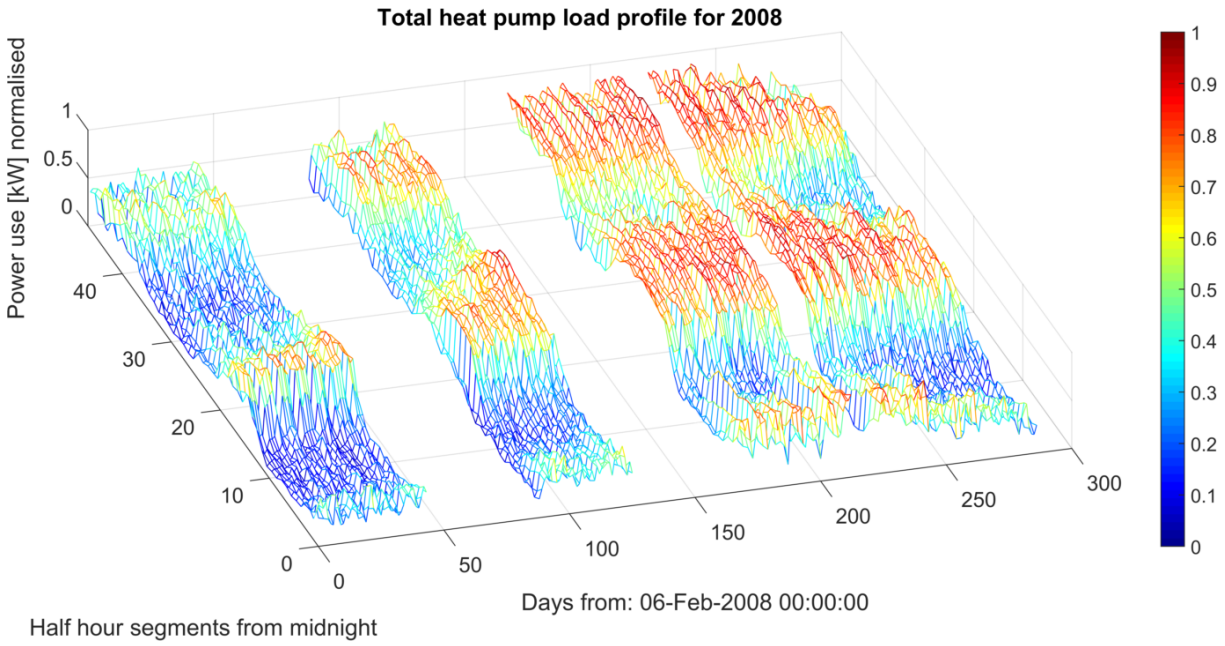

Figure 3-16: 3D plot of the normalised combined heat pump power profile for 2008.

3.7.5.3 Fold 2 load data for 2009

The heat map and 3D renderings of the combined heat pump power profile for 2009 is shown in Figure 3-17 and Figure 3-18 respectively.

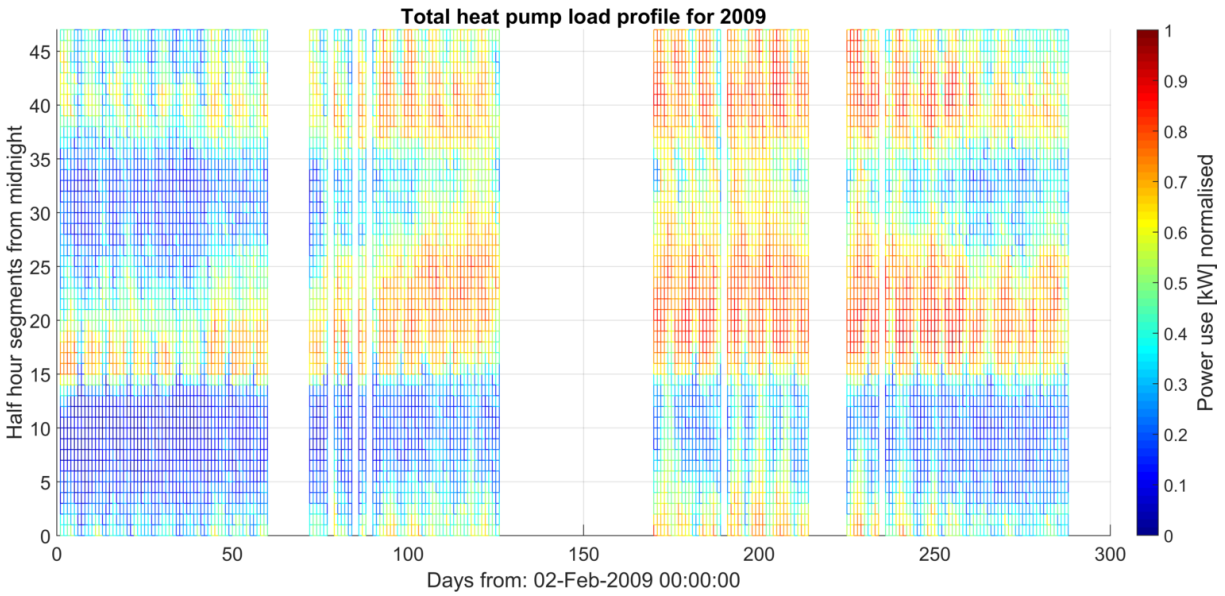

Figure 3-17: Heat map of the normalised combined heat pump power profile for 2009.

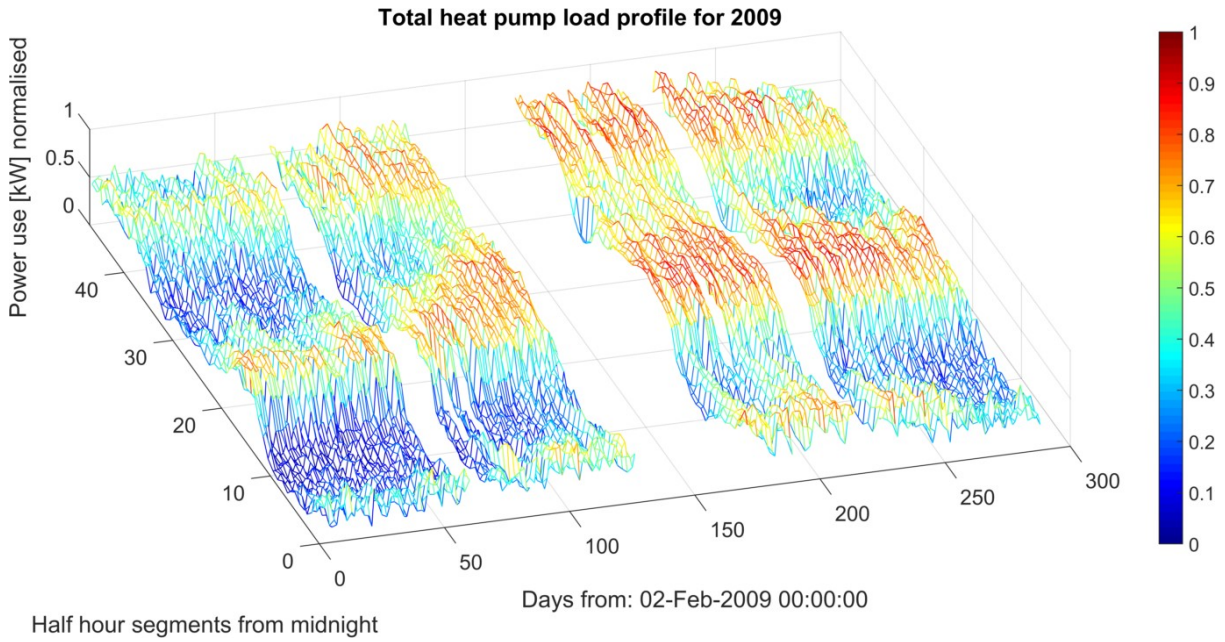

Figure 3-18: 3D plot of the normalised combined heat pump power profile for 2009.

3.7.5.4 Fold 3 load data for 2010

The heat map and 3D renderings of the combined heat pump power profile for 2010 is shown in Figure 3-19 and Figure 3-20 respectively.

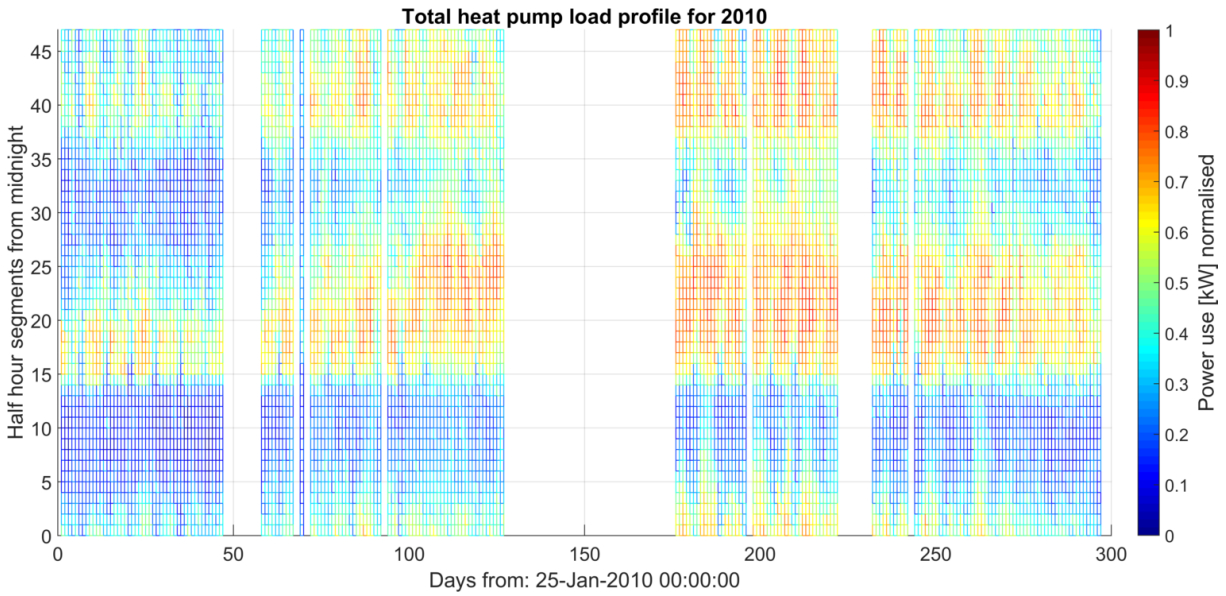

Figure 3-19: 2D plot of the normalised combined heat pump power profile for 2010.

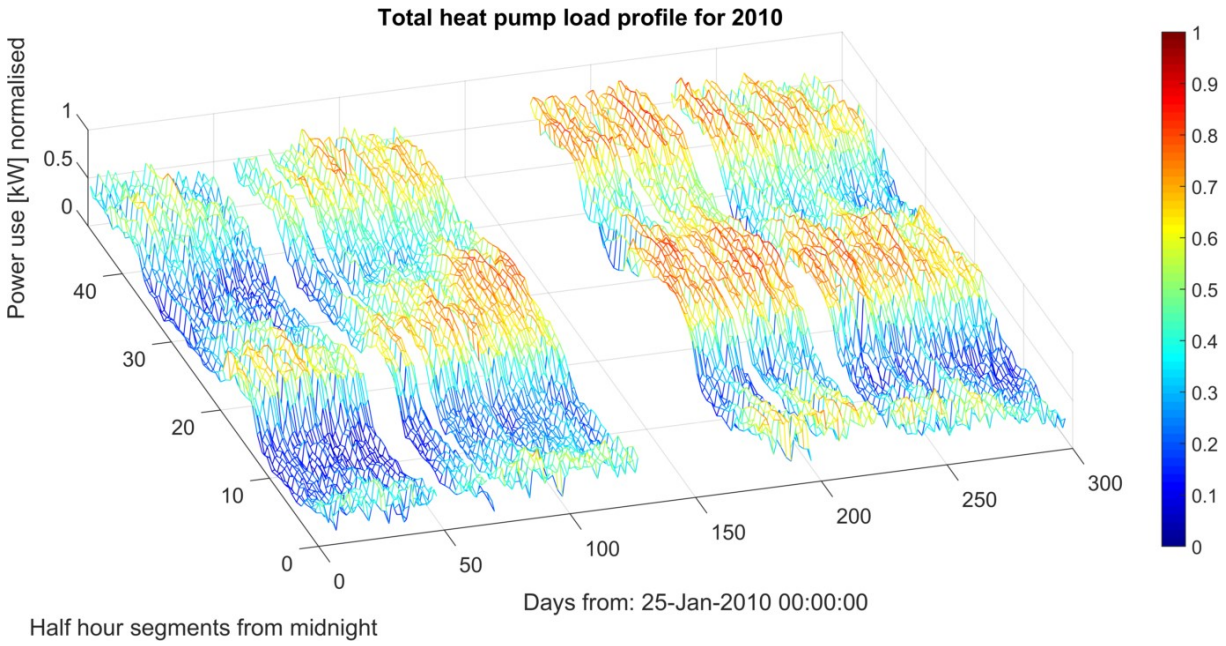

Figure 3-20: 3D plot of the normalised combined heat pump power profile for 2010.

3.7.5.5 Fold 4 load data for 2011

The heat map and 3D renderings of the combined heat pump power profile for 2011 is shown in Figure 3-21 and Figure 3-22 respectively.

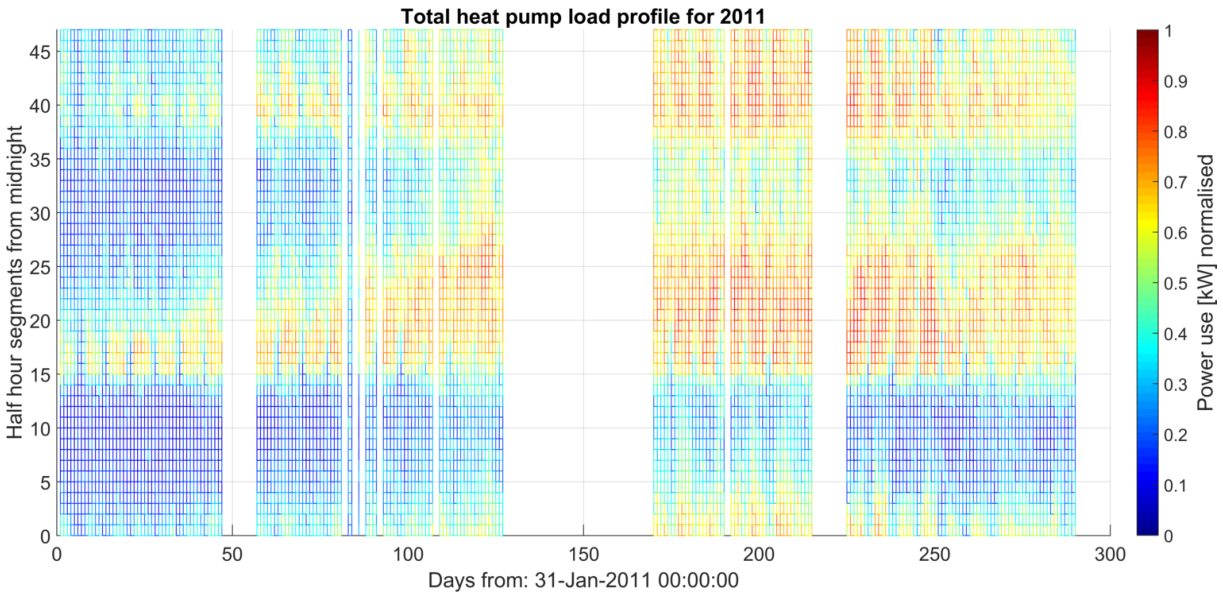

Figure 3-21: Heat map of the normalised combined heat pump power profile for 2011.

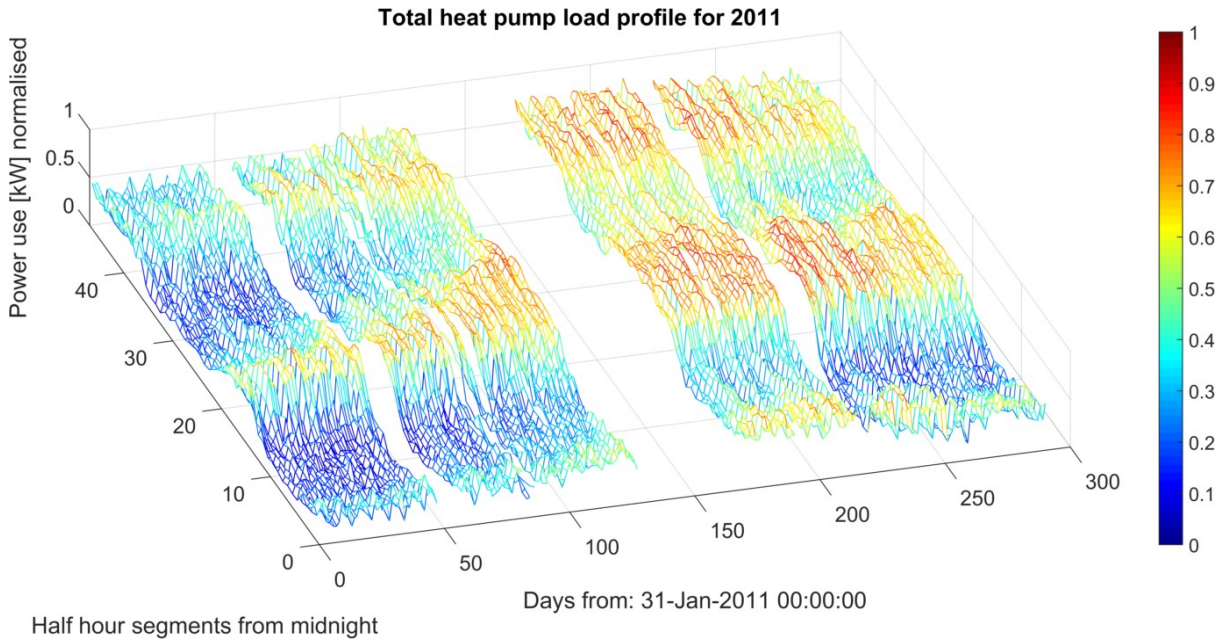

Figure 3-22: 3D plot of the normalised combined heat pump power profile for 2011.

3.7.6 Design choices

3.7.6.1 Initial FIS constructor routine

The three possible FIS constructor routines, *genfis1*, *genfis2* and *genfis3*, are described in section 2.5.3. The choice of routine depends on the preferred level of customisability of model features, the compatibility of the routine and the training data, and the intended purpose of the constructed FIS. The application of each is considered below, rated on these merits:

- *Genfis1* generates a single output Sugeno-type FIS using a grid partition on the data which is used as an initial state during ANFIS training. This routine is the only one out of the three that allows the user to explicitly choose the number of membership functions (MFs) for each input. The user can also choose an MF type other than the default Gaussian for every input. This level of customisability would allow for extensive experimentation with regards to the handling of the proposed model inputs. It was found that the drawback of this routine is that it produces a computationally cumbersome FIS with a large rule base, which grows exponentially with an increase in inputs as well as any increases in the size of the dataset used for training.
- *Genfis2* also generates a single output Sugeno-type FIS though subtractive clustering of the input space. The use of clustering instead of the grid partition of *genfis1* results in a smaller and more optimised set of fuzzy rules. The ideal numbers of MFs for each input is automatically determined, and are all of the type Gaussian.
- *Genfis3* only differs from *genfis2* because it uses fuzzy c-means clustering instead of subtractive clustering. It can also be used to produce a Mamdani-type FIS, which cannot be further tuned with ANFIS training.

The size of the load dataset alone disqualifies the *genfis1* routine as a viable option due to computational constraints. There have been successful implementations thereof for short term load forecasting, but all for short prediction horizons and training data spanning months, not years [11], [12]. The additional functionality of *genfis3* is not applicable to ANFIS training and *genfis2* is therefore used for the scope of all experimental models.

3.7.6.2 Number of training epochs

Each model is trained for 800 epochs to allow for enough passes of the back propagation algorithm to ensure a small RMSE value along with a training weight step size that does not oscillate heavily. Training over more epochs could result in slightly lower RSME values in some of the cases that are to follow, but to train for more would be excessive since the resulting improvement in accuracy is minimal when compared to increase in computational time.

Chapter 4: Software Implementation

4.1 Overview

This section documents the design of the software components developed during the course of the research.

4.2 Data formatting

4.2.1 Formatting objective

All of the proposed models are constructed within the framework of the MATLAB scripting language and platform. Therefore the objective of data formatting in this case is converting the obtained source data from its original file format to one that is directly compatible and readable by MATLAB. In this case that means conversion to m-files, which stores data in the matrix format native to MATLAB. The complete formatting procedure followed is documented further in this section, for both load and temperature data.

4.2.2 Formatting load data

The sourced load data for every heat pump has to be exported from the online Powerwatch server one month at a time to avoid lengthy server side compression operations, as that may hinder the performance of other applications thereon. The Data is received as a compressed monthly comma separated value (CSV) files containing timestamped average active- and reactive power readings. The formatting procedure followed with the load data is visualised in Figure 4-1, and is as follows:

- A Matlab script loads the monthly CSVs of each heat pump from a defined search path, compiles the data of each heat pump into a single file and saves the total range of load data for each heat pump as a CSV.
- The CSV file of each heat pump is imported into a load database architecture using the MySQL Query Browser interface.
- The active power data of each heat pump is queried from within the Matlab interface, through use of an ODBC driver. An intersect operation is run on the data that adds the load values of each unique timestamp together and removes the timestamp entries that are not present in all of the profiles.

- A two column m-file is exported that contains the cumulative heat pump load data from 2008 to 2011. The first column contains consecutive half-hourly forward filled timestamp values and the second contains the corresponding average power readings in kW per half hour.

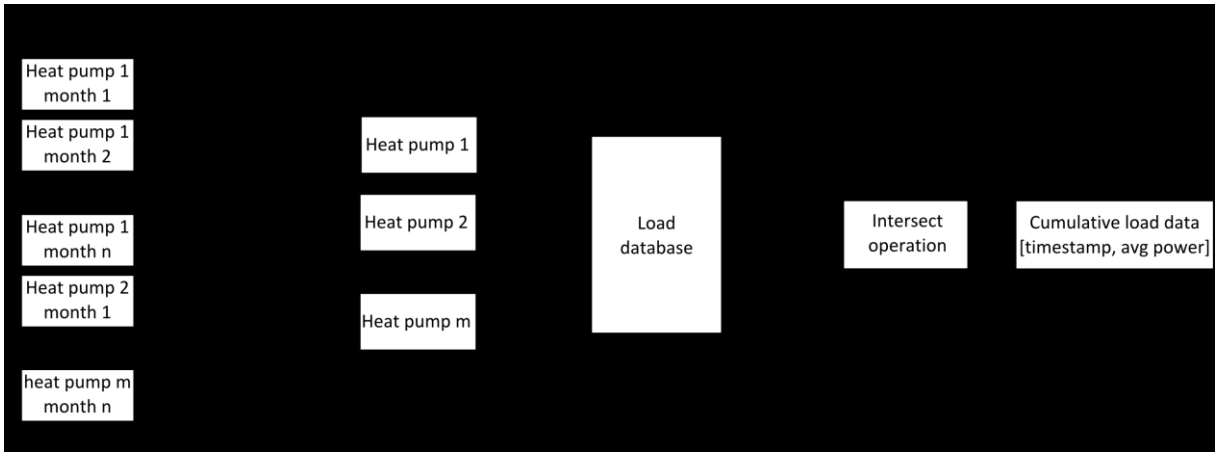

Figure 4-1: Data flow and operations for the formatting of the heat pump load data.

Saving the individual heat pump profiles in a database is a superfluous step for the purpose of this research. However it is still done as part of an on-going effort within the departmental research group to standardise the storage architecture of the wide range of load data profiles encountered. The reasons for this aspiration are ease of adoptability of the data from one project to the methods of another, which thereby assists in the productivity of future research.

4.2.3 Formatting of temperature data

The ambient temperature readings from the Cape Town International Airport weather station was received as a single CSV file. Unlike the load data the resolution of the temperature data was hourly and the values had no timestamp. Rather the values were presented as a grid table with rows that represent days of consecutive months and columns that represent each hour of the day. The formatting procedure is visualised in Figure 4-2 and is as follows:

- A Matlab script loads the CSV and performs the following operations:
  - Generates a timestamp value for each point of metered temperature data through its position in the CSV.
  - Expands the load data to a half-hourly resolution through a four point moving regression operation on the hourly data and assigns the proper timestamps to the new data.

- A two column m-file is exported that contains the temperature data from 2008 to 2011. The first column contains consecutive half-hourly forward filled timestamp values and the second contains the temperature in degrees Celsius.

Figure 4-2: Data flow and operations for the formatting of the temperature data.

4.2.4 Data sanitation

A data sanitation operation is performed on the contents of the load and temperature m-files by a Matlab script. The script generates a list of half-hourly timestamp values for the range of the data, and performs an intersection operation on the timestamps of the metered data to give a profile of full timestamps with nil values where temperature or load readings are missing. These missing values are filled in using a regression operation for interpolation. Had there been large gaps in the data, the period would have been removed as a pre-processing precaution.

4.2.5 Load data pre-processing

The data from the time periods in Table 3-2, Table 3-3 and Table 3-4 is removed from the load data file. A script loads the load data file and a file containing the dates of unwanted data. A list of the half hourly timestamps that would occur in the unwanted time periods is generated and the load data with corresponding timestamps is removed. The pre-processed load data is exported for later use.

4.3 Training and testing module

4.3.1 Data flow

Outlined in Figure 4-3 is the operation of a custom Matlab class used to train, store and validate multiple load models with differing configurations. A class is used so that a single class variable can store multiple load models, as well as the validation outputs from those models. The operation of the class is a follows:

- The class accepts as inputs time-stamped load and temperature data spanning several years, as well as the year that will be used for validation.

- The input data is converted from its time-stamped state to model input format. The process is shown in section 4.3.2.
- An operation designates the data from the chosen year into a validation dataset and creates a training dataset from the remaining data. These sets are stored as class variables.
- Different configurations of ANFIS load models are generated, trained and stored in a structured cell array. The different configurations are shown in section 4.3.3.
- A load profile is generated by each load model and stored in the class, using the inputs from the validation data.
- The output from the validation data, i.e. the load values, is stored as a class variable.

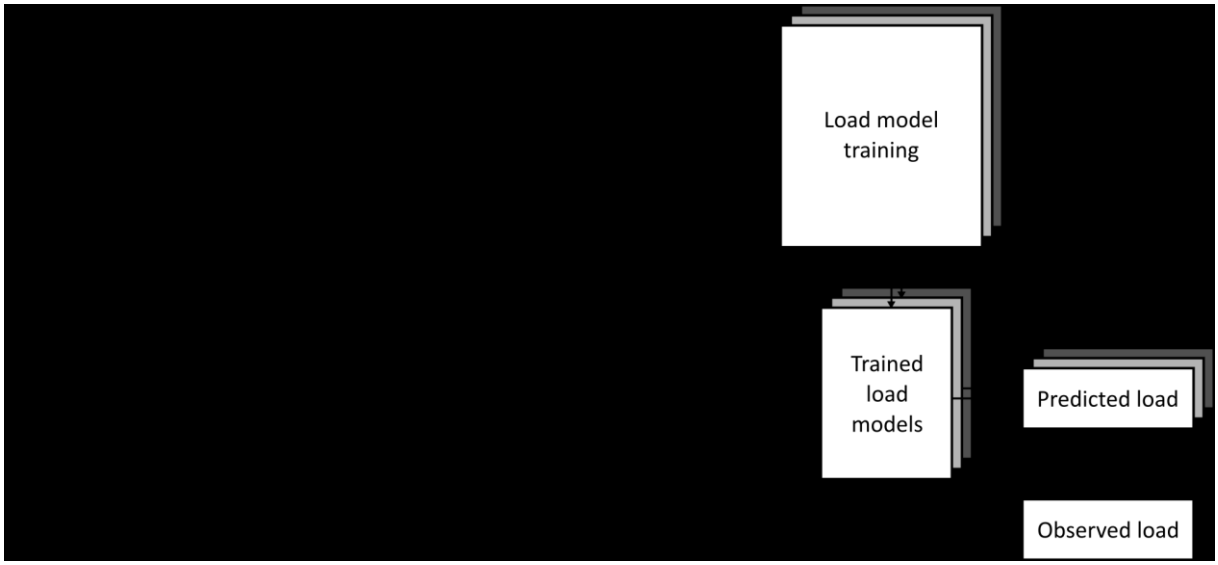

Figure 4-3: Data flow for model training and validation.

Two variants of this class are used for some of the load model experiments. The first of these variants does not accept a temperature input and therefore the constructor does not pass a temperature training input to the set of load model configurations. The second does not pass the day of year input. In both cases the load models are adapted to have one less input.

4.3.2 Dataset input constructor operations

The block diagram of this operation is shown in Figure 4-4. It converts the input data to suitable model input sets in the following steps:

- As the load data has been pre-processed to not include certain dates, so an intersection operation combines the load data with the appropriate temperature values.

- The day of year, day of week and timeslot inputs are calculated from the timestamp of each point of data.
- The set of data for each input is normalised to a unity scale, as ANFIS models require normalised inputs.
- Each input set is an array of consecutive values and the position of each value in the array aligns with a complementary timestamp value in the same position in another array. In the instances where the input data is sub-divided, this timestamp value accompanies all of the data so that the profile can later be reconstructed through sorting of this timestamp index.

Figure 4-4: Operations of input constructor.

4.3.3 Training and experimental ANFIS configurations

4.3.3.1 Single ANFIS

A single ANFIS is trained with all of the training inputs and the resulting model is saved. The configuration of this model is shown in Figure 4-5.

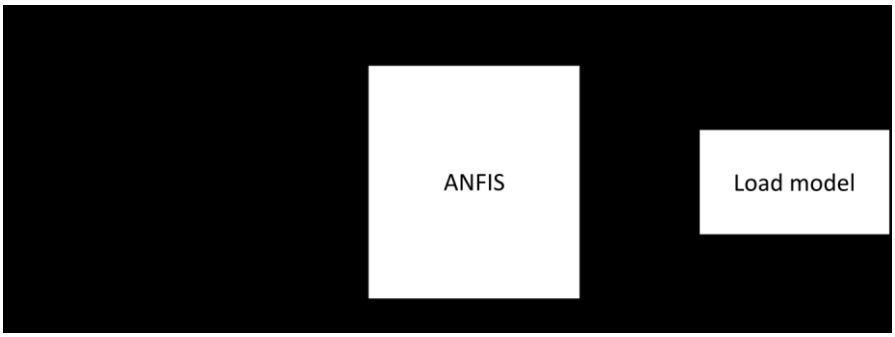

Figure 4-5: Single ANFIS training configuration.

4.3.3.2 Weekday category sub-model ANFIS

This load model splits the data into three sub-sets, each targeting a different weekday category. As there is an ANFIS for each weekday category, the day of week input set is not passed as an input. Instead it is used by an indexing operation to sort the remaining training inputs into the correct day of week category. The sub-model configuration as well as the

indexing operation is visualised in Figure 4-6. The trained ANFIS of each day of week category is saved separately.

Figure 4-6: Weekday category sub-model ANFIS configuration and operations.

4.3.3.3 Time of day sub-model ANFIS

This load model splits the data into 48 sub-sets, each targeting one of the 48 half-hourly intervals of a day. As there is an ANFIS for each time of the day category, the timeslot input set is not passed as an input. Instead it is used by an indexing operation to sort the remaining training inputs into the correct time of day category. The operational block diagram is shown in Figure 4-7. The trained ANFIS of each time of day category is saved separately.

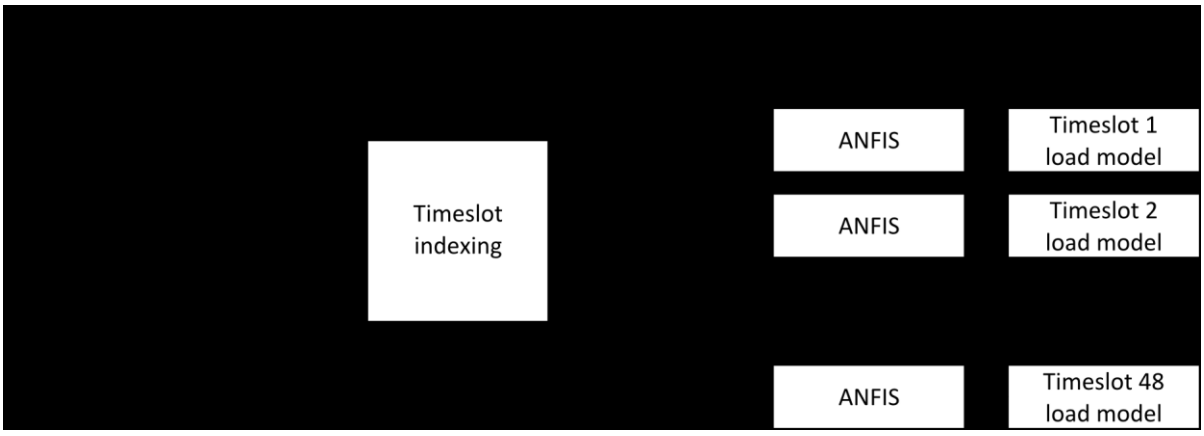

Figure 4-7: Time of day sub-model ANFIS configuration and operations.

4.3.3.4 Time of day in weekday category sub-model ANFIS

This load model is an amalgamation of the previous two, as there is a distinct ANFIS for every time of day category within each weekday category. The timeslot and day of week input sets are not passed as training inputs, but are used to sort the training inputs into the correct categories. The operational block diagram is shown in Figure 4-8. The trained distinct ANFIS of every combination of time of day and weekday category are stored separately.

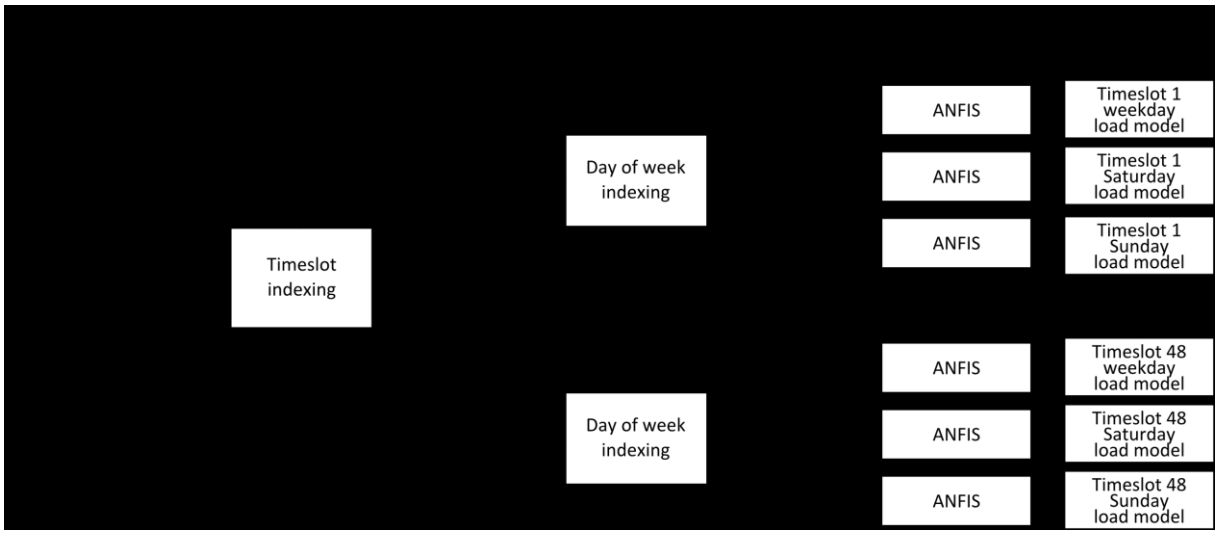

Figure 4-8: Time of day in weekday category sub-model ANFIS configuration and operations.

4.3.3.5 Class variants

As mentioned before, there are two variants of the training and validation class. One that does not utilise the temperature input set, and one that does not utilise the day of year input set. The first three of the abovementioned load models are adapted into these variant classes, simply by removing the respective input sets from the group of training inputs. Removing either the temperature or day of year input from fourth load model configuration would result in an ANFIS with only one input. Therefore that load model cannot be implemented by these class variants.

4.3.3.6 Validation load model output

Directly after training and storing each ANFIS model or sub-model, the model is tested with validation data and the output stored. For the cases where the load model is sub-divided the validation data is indexed and manipulated in the same manner as the training data, with the use of indexing terms. The outputs of each sub-model within a load model are recombined using the timestamp variable.

4.4 Accuracy report module

The training inputs, testing inputs, model files and model outputs for each cross validation fold is stored in training and validation class files. A second class was created that accepts one of these trained class files as input, then performs a comprehensive accuracy analysis on the load models stored therein and stores the accuracy results within a sub-divided cell array architecture. The accuracy analysis involves the comparison of the various load model outputs

with the observed load through error calculations. The root mean squared error (RMSE), percentage error (PE) and Mean Absolute Percentage Error (MAPE) accuracy measures are used to calculate the time-series accuracy of the load model outputs. The time-series accuracies of each of these methods is calculated and stored for the total output, outputs of specific weekday categories and outputs of specific time of day categories. Lastly the percentage error of the total energy of the observed load and the total energy of the load model is calculated and stored.

The accuracy class contains custom graphing methods that can generate the following mesh plots:

- Plot of the temperature over any year of the training or testing data.
- Plot of the observed load over any year in the training or testing data.
- Plot of the output of a load model.

Chapter 5: Modelling Results and Performance Evaluation

5.1 Overview

This chapter contains the following:

- A description of the objectives of the case study.
- A guide to the methodology followed during evaluation of each load model.
- An exposition and analysis of the predicted load profiles and prediction accuracies of each experimental load model.
- A summary of the relative performances of the load models.
- Observations on the relative computational complexity of the load models.
- An analysis of the time-series prediction accuracies.
- An analysis of the annual energy prediction accuracies.

5.2 Case study objectives

The objectives of the case study are as follows:

- Find an ANFIS configuration that can adequately model residential load behaviour in the medium term given certain descriptive inputs.
- Compare the strengths and weaknesses of different ANFIS model configurations through:
  - Permutations of differing model inputs.
  - Reducing the complexity of the models by breaking each model into various sub models that are each trained with a different sub-set of the data.

In this case the former objective can be satisfied through completion of the latter, as the best performing model can be distinguished from the case study pool.

5.3 Case study methodology

5.3.1 Overview

The case study methodology section consists of a summary table containing the variable parameters for each case, followed by the presentation steps deemed necessary for each experimental load model in order to ensure model validity and relative accuracy. These steps are:

- Graph the output of each model for each cross validation fold.
- Compile a comprehensive accuracy report on the performance of each model.
- Test for over fitting of the model to the testing data.

5.3.2 Case study parameter summary

The chosen possible training inputs have been discussed at length, as well as the reasoning and procedure of dividing models into sub-model systems. The model configuration of each case study is tabulated below in Table 5-1.

Table 5-1: Model configurations for case studies.

	Inputs				**Sub-models per:**		**Models**
Case	**Temperature**	**Day of year**	**Day of week**	**Time of day**	**Weekday category**	**Timeslot**	
A1	✓	✓	✓	✓			$A1_1$
A2	✓	✓		✓	✓		$A2_1 – A2_3$
A3	✓	✓	✓			✓	$A3_1 – A3_{48}$
A4	✓	✓			✓	✓	$A4^1_1 – A4^1_{48}$ $A4^2_1 – A4^2_{48}$ $A4^3_1 – A4^3_{48}$
B1		✓	✓	✓			$B1_1$
B2		✓		✓	✓		$B\ 2_1 – B\ 2_3$
B3		✓	✓			✓	$B\ 3_1 – B\ 3_{48}$
C1	✓		✓	✓			$C1_1$
C2	✓			✓	✓		$C2_1 – C2_3$
C3	✓		✓			✓	$C3_1 – C3_{48}$
D1	✓ (Delayed)	✓	✓	✓			$D1_1$
D2	✓ (Delayed)	✓		✓	✓		$D2_1 – D2_3$
D3	✓ (Delayed)	✓	✓			✓	$D3_1 – D3_{48}$
D4	✓ (Delayed)	✓			✓	✓	$D4^1_1 – D4^1_{48}$ $D4^2_1 – D4^2_{48}$ $D4^3_1 – D4^3_{48}$

5.3.3 Graph model outputs

The performances of the various load model designs are investigated with statistical analysis and accuracy calculations, but a simple visual inspection of model outputs can lend valuable insights into behaviours that might go unnoticed otherwise. Therefore the predicted model

outputs for each cross-validation fold are generated and displayed as a heat-map rendering, and may serve as comparison to the observed training data profiles in section 3.7.5.

5.3.4 Performance evaluation

In order to determine the weaknesses of each model, and to enable comparison between aspects of different models, a comprehensive accuracy report is generated for each model. In the same manner that the models are compartmentalised into sub models, the accuracy reports are designed to show the compartmentalised accuracies of the following sub sections of data:

- The total time-series accuracy.
- The accuracies realised in each weekday category.
- The accuracies realised in each timeslot category.

Additional to the abovementioned accuracy details of the model outputs as a time-series of load power use values, the total energy consumption of the observed load is compared with the total energy consumptions predicted by the models. This is done by converting the half-hourly power data to half-hourly energy so the total energy used within the specified time ranges can be calculated.

5.3.5 Test for over-fitting

When a model is trained using ANFIS, the initial FIS is constructed and thereafter undergoes changes to its node function parameters as well as input and output membership functions as the training routine is executed for a number of iterations. Each of these iterations of training is called an epoch. Training and checking error can be obtained for each epoch, where the training error is the RMSE of the model when tested with the same data it was trained with, and the checking error is the RMSE of the model when tested with data not included in the training set. A lot can be told of the validity of a model by observing what these training and checking errors are, and how they change with relation to each other over the range of epochs. It is generally expected that the training error will be lower than the checking error, as the model is presented with data that is unseen during training when the checking error is obtained. If the model were to perform very poorly with the checking data, but well with the training data, it would be evidence that the model had been over-fit to the training data [40].

In order to avoid this, the training and checking errors are saved and graphed over the epoch range for each model or sub-model component, so that any cases of over-fitting would be

easily identifiable. As an example of this, Figure 5-1 shows the training and checking RMSE of model A1 for each of the four folds (specified in Table 3-8) in the k-fold cross validation method, graphed over the training epochs. For all of the folds the training and checking errors both decrease evenly as the model parameters are adjusted and become better suited to the target data behaviour, and at no point does the checking error increase. For each fold the difference between checking and training errors differ, and for 2011 the checking error is even consistently lower than the training error, though this can be ascribed to the decrease in the demand of the load over the range of training years, as discussed in section 3.6.2.

Figure 5-1: Training and checking error RMSE values over course of training epochs for each checking year in the k-fold cross validation method.

5.4 Group A case study

5.4.1 Model A1

5.4.1.1 Overview

Model set A1 consists of a single model targeting any day of the year, any day of the week and any time period of the day. The model accepts all inputs, i.e. temperature, day of year, day of week and time of day. The entire data set is used for training and testing.

5.4.1.2 Model outputs

Shown below is the output profile of model A1 for fold 4 of cross validation training, i.e. the prediction for 2011 when trained with data from the years 2008 to 2010. Figure 5-2 is a heat

map rendering of the output for 2011, and Figure 5-3 is the 3D mesh rendering for the same output data. The model outputs for fold 1 to 3 can be found Appendix A.1. Upon inspection of the model prediction outputs, it can be seen that model A1 clearly succeeds in presenting a load profile with characteristics similar to the observed data (shown in Figure 5-23 and Figure 5-24). Throughout the year the profile exhibits low power use between midnight and roughly 7:30 AM (timeslot 15), there are two periods of peak usage throughout the day, and there is much higher power use in the second half of the year than the first. It is however clear from Figure 5-3 that the output profile of model A1 shows gradual increases and decreases in the daily power use, unlike the observed data, which exhibits a high ramp rate in the load over the interval between 7:00 AM and 8:00 AM. This gradual change in load behaviour throughout the day gives the profile surface a smooth appearance.

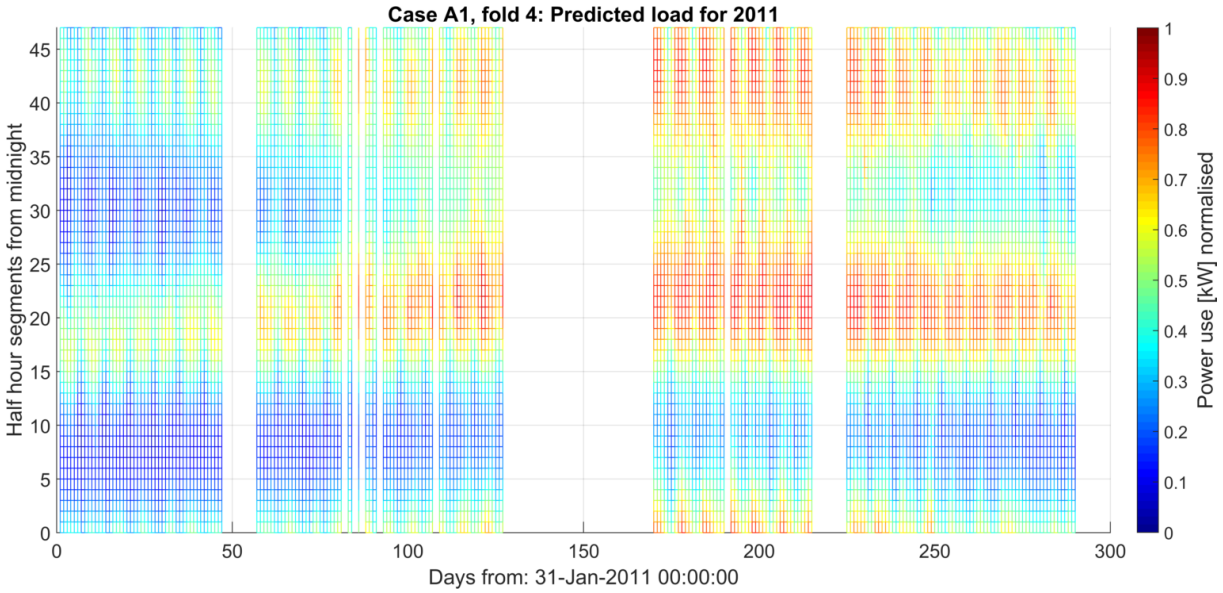

Figure 5-2: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model A1.

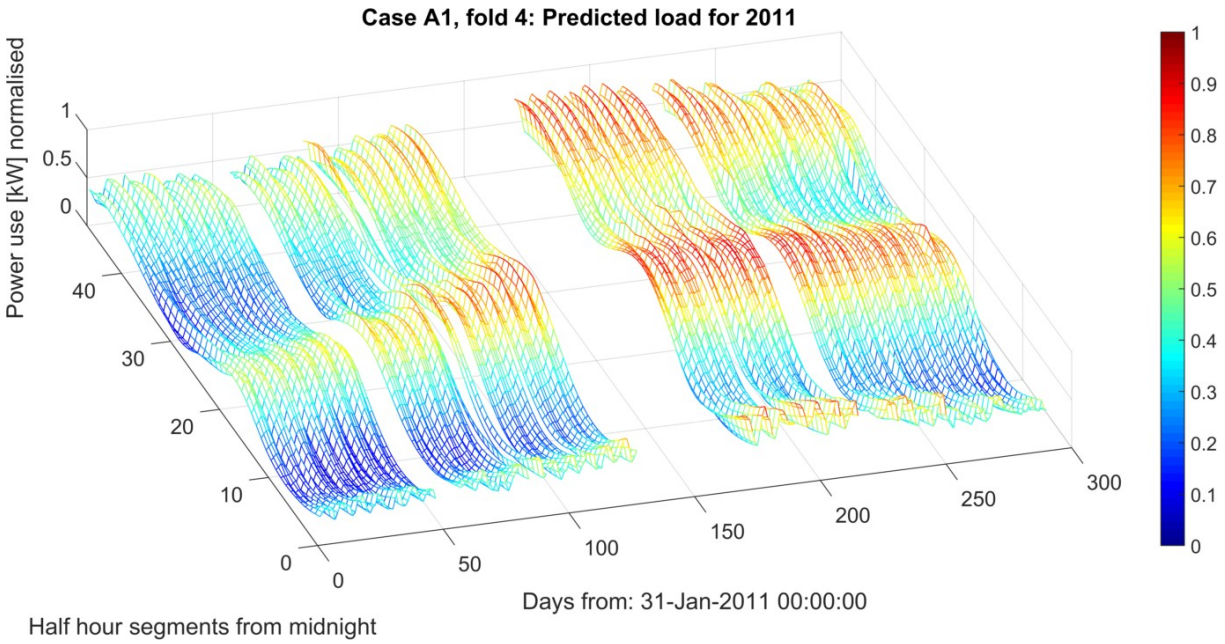

Figure 5-3: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model A1.

5.4.1.3 Comprehensive accuracy assessment

The accuracy measure of each cross validation fold is shown in Table 5-2. The MAPE achieved by model A1 in each fold is shown, as well as the MAPE of each weekday category within each fold. The mean of the error of all folds is also shown for each weekday category. Model A1 has a time-series prediction MAPE of 24.6%, the error is slightly higher for prediction of weekends and the poorest performance occurred for the prediction of 2009. The MAPE of each fold is graphed over all timeslots in Figure 5-4. From this it is apparent that the model exhibits a much higher rate of error for predictions in the intervals of timeslot 3 to 16, which coincides roughly with the periods of reduced power use associated with the early mornings. The model seems to perform especially poorly during this period when predicting 2009, attributing to the high total MAPE for that year. There is a second period of increased error rate in the evening between timeslot 30 and 37, although far less pronounced, which coincides with the traditional periods of reduced power use seen between the morning and evening peak.

Table 5-2: MAPE of each cross validation fold over each weekday category using model A1.

Fold	Testing year	MAPE (%)			
		Whole week	Weekdays	Saturdays	Sundays
1	2008	24.72	24.50	25.56	24.96
2	2009	26.63	26.39	28.21	26.27
3	2010	25.06	24.48	26.23	26.94

4	2011	22	21.96	20.91	23.30
Mean		24.60	24.33	25.23	25.37

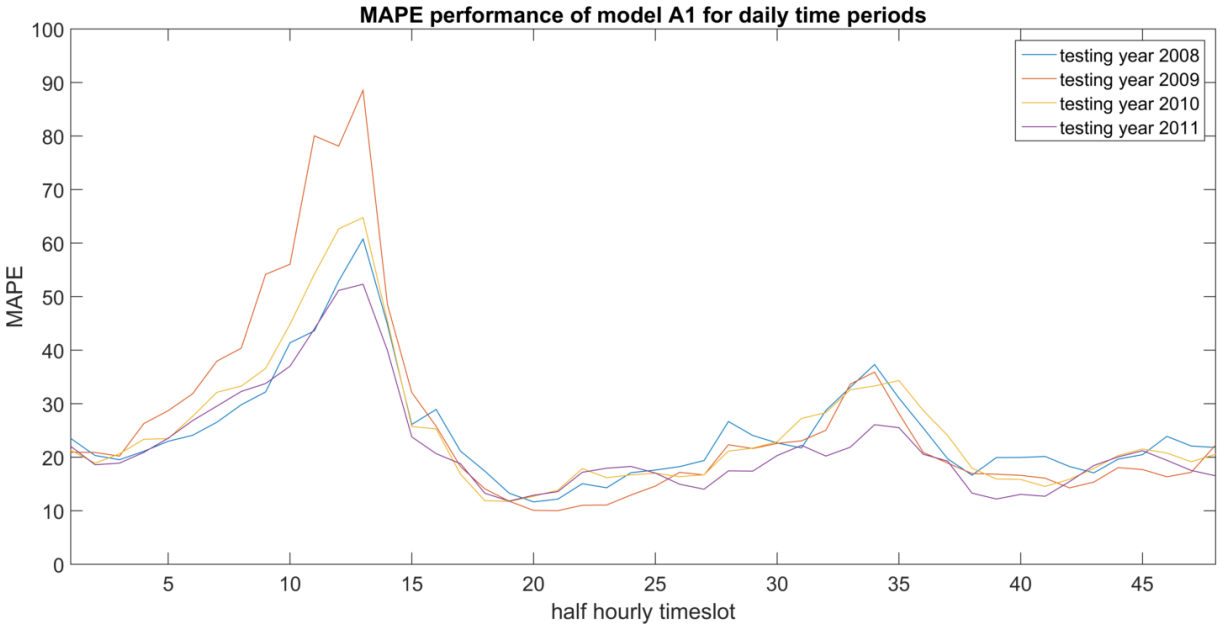

Figure 5-4: MAPE over each cross validation fold over each daily timeslot using model A1.

The accuracy obtained in each fold for predicting the total energy consumption for the year long period is shown in Table 5-3. Take note that the total energy refers only to the total of energy used within the timeframes not removed during pre-processing. The accuracy of the model at this resolution is much higher than that obtained for the time-series accuracy, which indicates that the under predictions compensate for the over predictions to some extent in some cases, and vice-versa for others. A percentage error of 2.91% shows that the model under-predicted by that percentage for 2008 and 2009, and over predicted for 2010 and 2011. The mean energy consumption of the observed data is higher for 2008 and 2009 and lower in 2010 and 2011. This could attribute partially to the percentage errors obtained.

Table 5-3: Percentage error of model A1 in estimating yearly energy consumption within the training intervals.

Year	**Total observed energy consumption [MWh]**	**Total predicted energy consumption [MWh]**	**Percentage error (%)**	**Over/under predicted**
2008	670.8	651.25	2.91	Under
2009	739.14	717.59	2.91	Under
2010	727.01	746.74	-2.71	Over

2011	734.26	755.73	-2.92	Over

5.4.2 Model A2

5.4.2.1 Overview

Model set A2 consists of 3 models targeting weekdays, Saturdays and Sundays respectively. The models accept temperature, day of year and time of day as inputs. The data sets used for training and testing the models consist of subsets representing weekdays, Saturdays and Sundays respectively.

5.4.2.2 Model outputs

The output profile of model A2 for fold 4 of cross validation training is shown as a heat-map in Figure 5-5 and as a 3D mesh in Figure 5-6. The outputs for fold 1 to 3 are shown in Appendix A.2. The profile has similar visual characteristics to that of model A1, with the notable difference that all of the days of a given week seem to have a very similar profile. This is because sub model in A2 that targets weekdays does not receive the day of the week as input. Therefore the model does not make distinction based on the specific day of the week and treats them relatively equal. Changes in the daily load profile are only derived from changes in the time of the year and the temperature, which changes slowly.

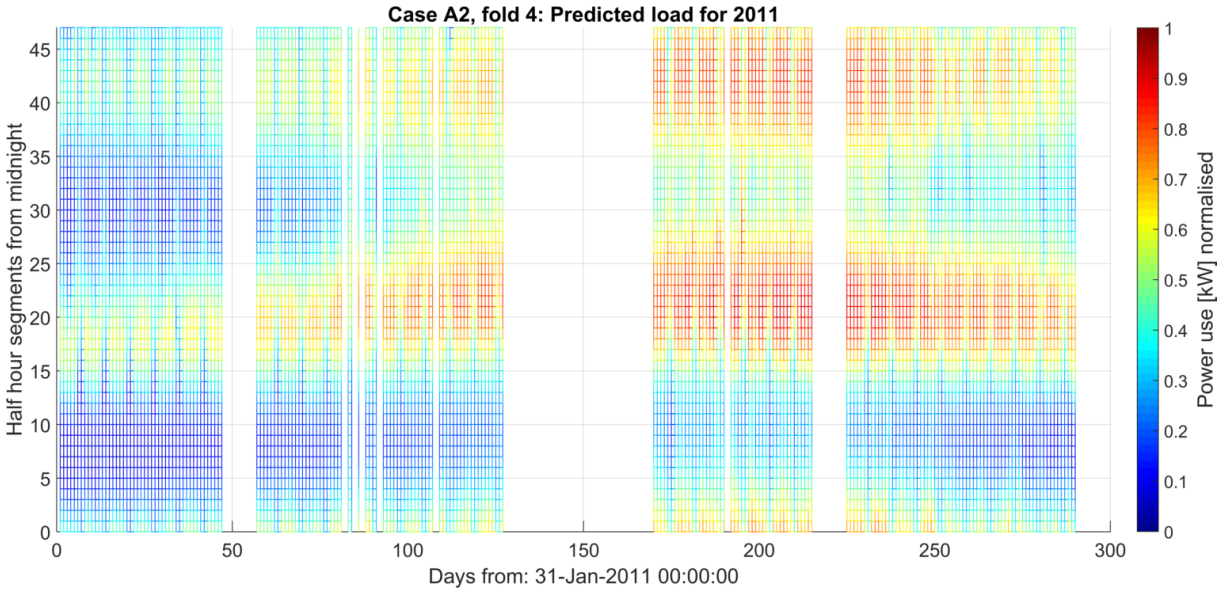

Figure 5-5: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model A2.

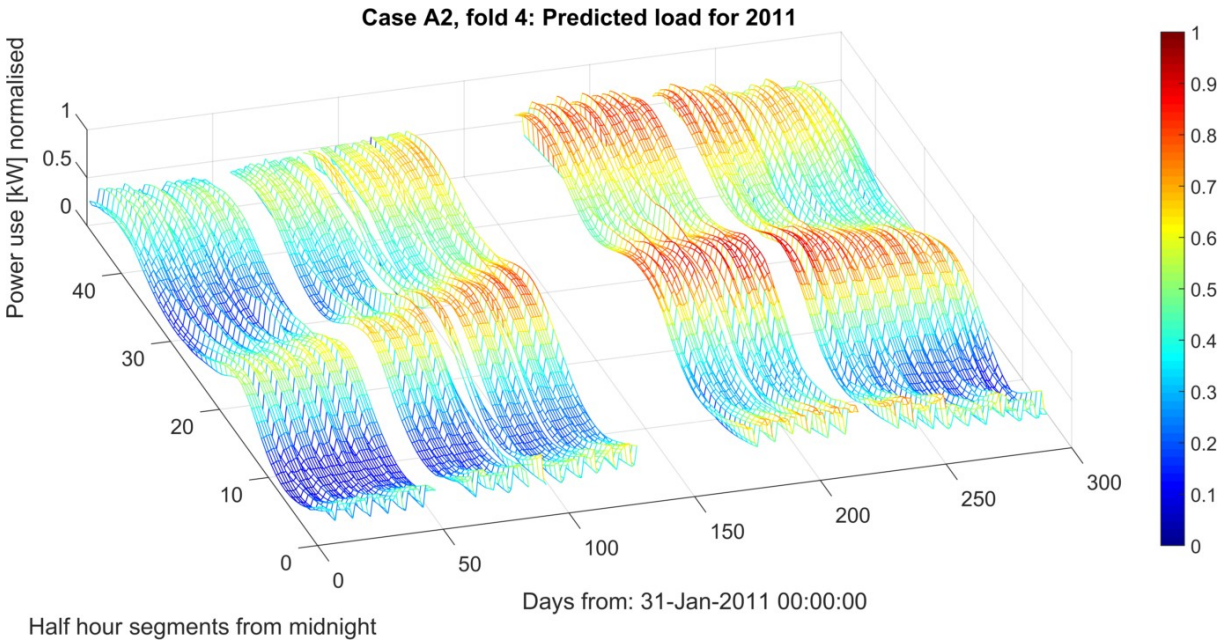

Figure 5-6: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model A2.

5.4.2.3 Comprehensive accuracy assessment

The accuracies of model A2 are shown in Table 5-4. The MAPE of each fold is only slightly lower than that of model A1, by less than 1%. The same can be seen when comparing the mean weekday category errors. Figure 5-7 shows the MAPE of model A2 for each fold over the range of daily timeslots, and the performance seems to be similar to that of A2.

Table 5-4: MAPE of each cross validation fold over each weekday category using model A2.

Fold	**Testing year**	**MAPE (%)**			
		Whole week	**Weekday**	**Saturdays**	**Sundays**
1	2008	24.16	23.65	25.36	25.57
2	2009	26.94	26.53	28.58	27.35
3	2010	24.02	23.63	25.48	24.59
4	2011	21.66	21.60	21.96	21.65
Mean		24.19	23.85	25.35	24.79

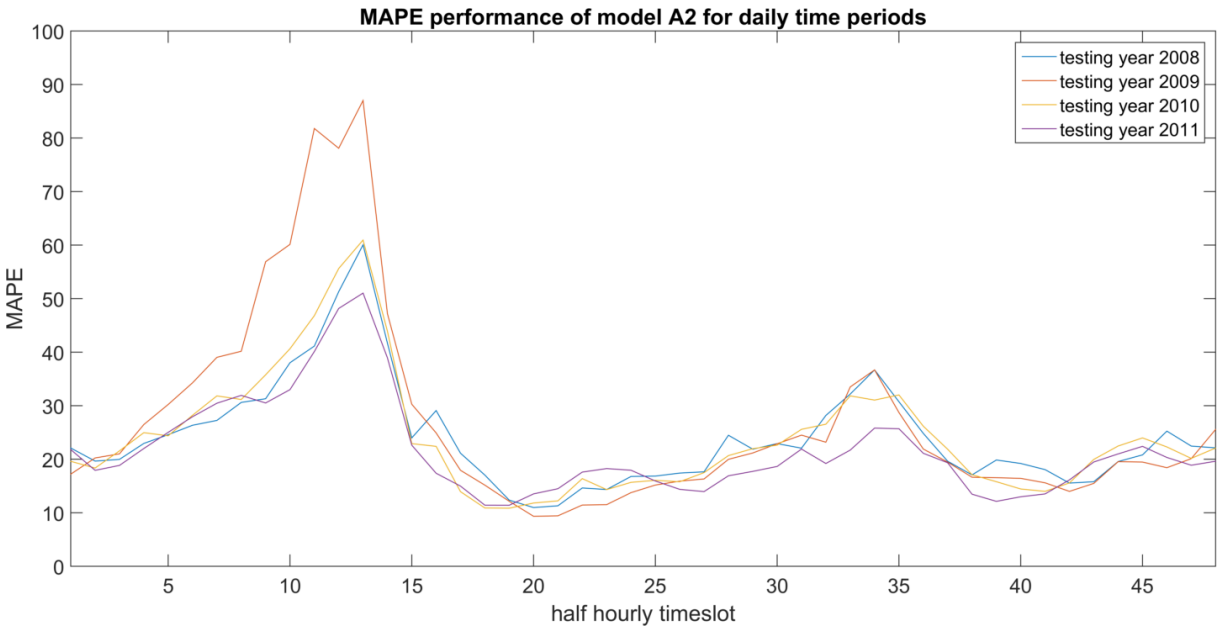

Figure 5-7: MAPE over each cross validation fold over each daily timeslot using model A2.

The accuracy of the total energy consumption in each fold is shown in Table 5-5. The Percentage error of each fold is similar to that of model A1, with over and under predictions in the same folds.

Table 5-5: Percentage error of model A2 in estimating yearly energy consumption within the training intervals.

Year	Total observed energy consumption [MWh]	Total predicted energy consumption [MWh]	Percentage error (%)	Over/under predicted
2008	670.8	649.62	3.16	Under
2009	739.14	717.68	2.9	Under
2010	727.01	746.83	-2.73	Over
2011	734.26	755	-2.82	Over

5.4.3 Model A3

5.4.3.1 Overview

Model set A3 consists of 48 models targeting the 48 daily time periods respectively. The models accept temperature, day of the year and day of the week as inputs. The data sets used for training and testing the models consists of subsets representing the 48 daily time periods respectively.

5.4.3.2 Model outputs

The output profile of model A3 for fold 4 of cross validation training is shown as a heat-map in Figure 5-8 and as a 3D mesh in Figure 5-9. The outputs for fold 1 to 3 are shown in Appendix A.3. The profile has similar visual characteristics to the output of model A1, but where the output profile of model A1 appears smooth, that of model A3 is not. This gives the predicted profile a more "real" look, which better resembles that of the observed data. The profile generated by model A3 seems to better match the rapid increases in power use that occurs every day at roughly 7:30. It would appear that by having a sub-model serve data from each timeslot individually, the complexity of the task of each model is reduced.

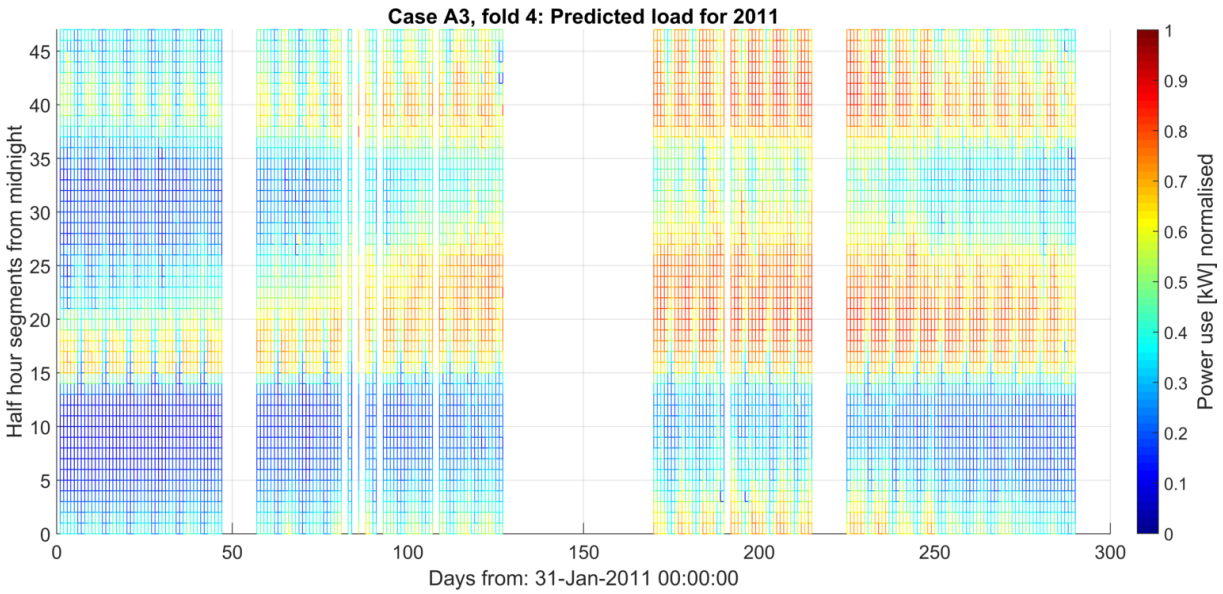

Figure 5-8: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model A3.

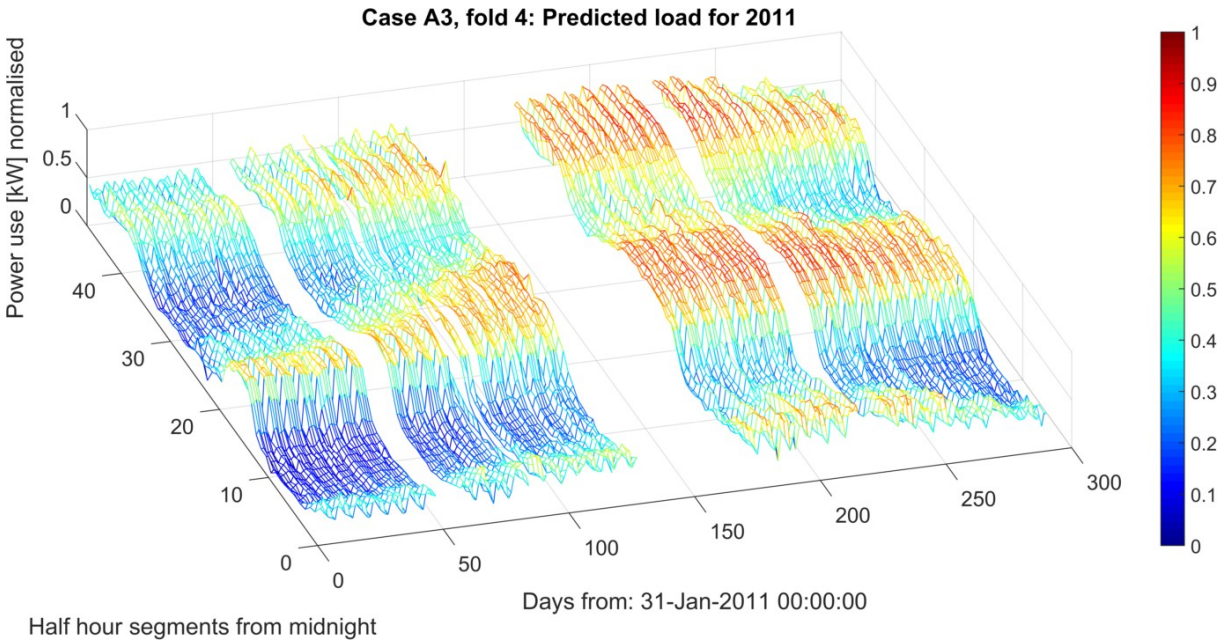

Figure 5-9: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model A3.

5.4.3.3 Comprehensive accuracy assessment

The accuracies of model A3 are shown in Table 5-6. With a mean MAPE of 21.02 % over all folds, this model out-performs models A1 and A2. The MAPE of weekdays are lower than that of Saturdays and Sundays, as there are far more weekdays with which the models can be trained than weekend days. Figure 5-10 shows the MAPE over each timeslot. As with models A1 and A2, the time periods between timeslot 3 and 16 show an increased error rate for that period, although to a lesser extent. The peak in the error profile later in the day is also reduced. The largest MAPE for this period still occurs for prediction of 2009, which once again is much worse than for the other years.

Table 5-6: MAPE of each cross validation fold over each weekday category using model A3.

Fold	Testing year	MAPE (%)			
		Whole week	Weekday	Saturdays	Sundays
1	2008	20.75	19.56	24.75	22.74
2	2009	22.94	21.96	26.50	24.34
3	2010	21.02	20.03	24.64	22.61
4	2011	19.37	18.58	21.16	21.72
Mean		21.02	20.03	24.26	22.85

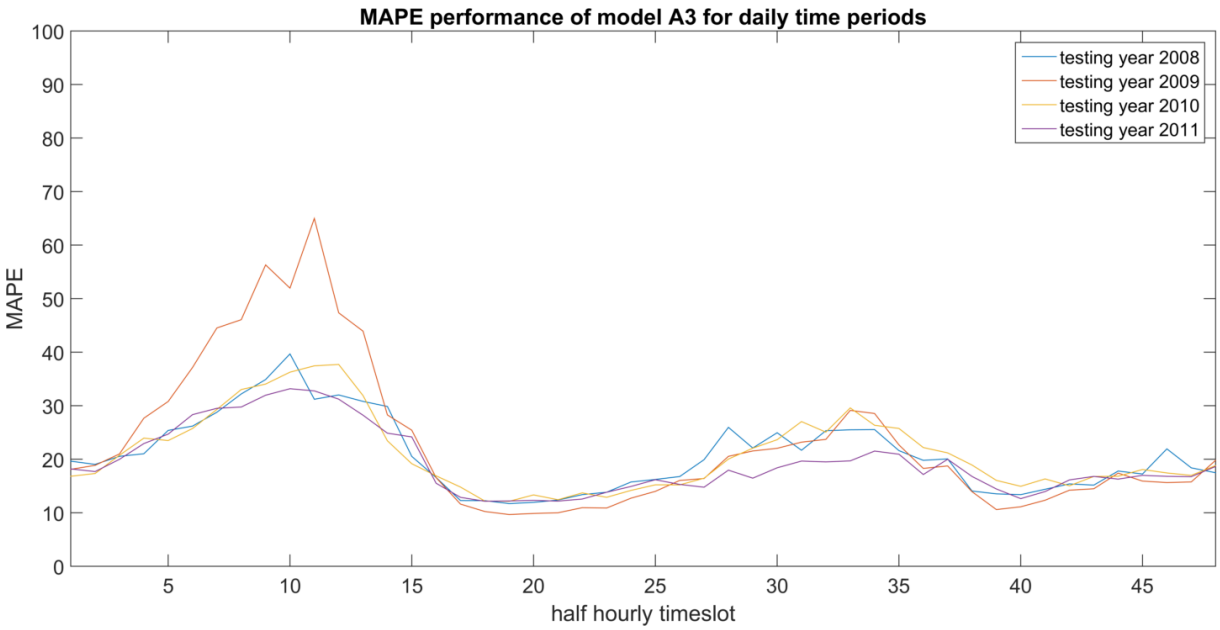

Figure 5-10: MAPE over each cross validation fold over each daily timeslot using model A3.

The accuracy of the total energy consumption in each fold is shown in Table 5-7. The increased time series accuracy of model A3 does not translate to better performance in predicting total energy consumption, as the percentage error for each fold is similar in orientation and magnitude to the values obtained in models A1 and A2.

Table 5-7: Percentage error of model A3 in estimating yearly energy consumption within the training intervals.

Year	Total observed energy consumption [MWh]	Total predicted energy consumption [MWh]	Percentage error (%)	Over/under predicted
2008	670.8	650.26	3.06	Under
2009	739.14	718.24	2.83	Under
2010	727.01	747.1	-2.76	Over
2011	734.26	754.87	-2.81	Over

5.4.4 Model A4

5.4.4.1 Overview

Model set A4 consists of a 3 subsets of models targeting weekdays, Saturdays and Sundays respectively. Each subset consists of 48 models targeting the 48 daily time periods respectively. The models accept temperature and day of the year as inputs. The data sets used for training and testing the models consists of subsets representing the 3 day of week categories and associated 48 daily time periods respectively.

5.4.4.2 Model outputs

The output profile of model A3 for fold 4 of cross validation training is shown as a heat-map in Figure 5-11 and as a 3D mesh in Figure 5-12. The outputs for fold 1 to 3 are shown in Appendix A.4. It is interesting that the profile generated by model A4 exhibits characteristics seen in both model A2 and A3, as it is a combination of the sub model division of both. It exhibits the similarity of the daily profiles over a week shown in model A2, and the improved load ramp rate of model A3. Model A2 has a sub-model for each weekday category, model A3 for each timeslot, and model A4 for each weekday category in each timeslot. However it is clear that there is something wrong with this model, as there are frequent large errors for weekend profiles that occur at random times. This erratic behaviour for weekend predictions is probably due to sub-dividing the data into too many smaller sets for increasingly specialised sub-models, resulting in far smaller training datasets for weekends than weekdays.

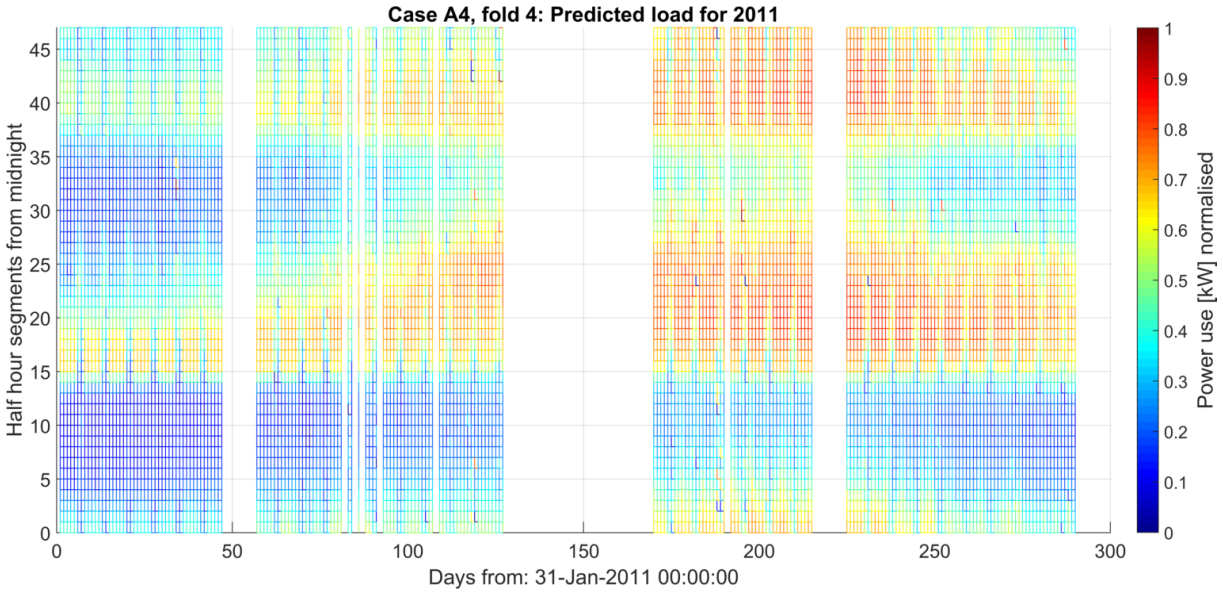

Figure 5-11: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model A4.

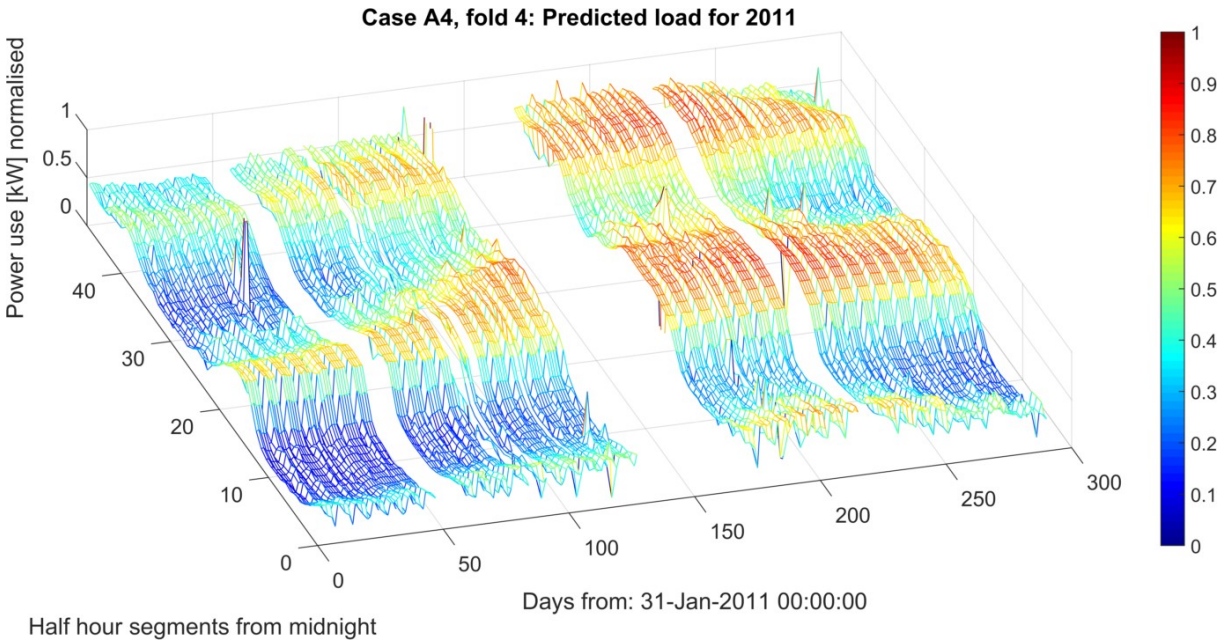

Figure 5-12: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model A4.

5.4.4.3 Comprehensive accuracy assessment

The accuracies of model A4 are shown in Table 5-8. The accuracies obtained for weekday predictions are similar to those of model A3, but the weekend categories have a much higher MAPE of 26.83 % for Saturday and 26.08 % for Sunday. Much higher than model A3, reflecting the effect of the outlier predictions of weekends on the total accuracy of the model. Figure 5-13 shows the MAPE over each timeslot. Behaviour over previously mentioned problem time periods is similar to that of model A3, which is better than models A1 and A2.

Table 5-8: MAPE of each cross validation fold over each weekday category using model A4.

Fold	Testing year	MAPE (%)			
		Whole week	Weekday	Saturdays	Sundays
1	2008	21.80	20	27.08	25.62
2	2009	23.75	22.10	28.96	26.89
3	2010	21.96	20.07	27.19	26.71
4	2011	20.31	18.68	24.08	25.09
Mean		21.96	20.21	26.83	26.08

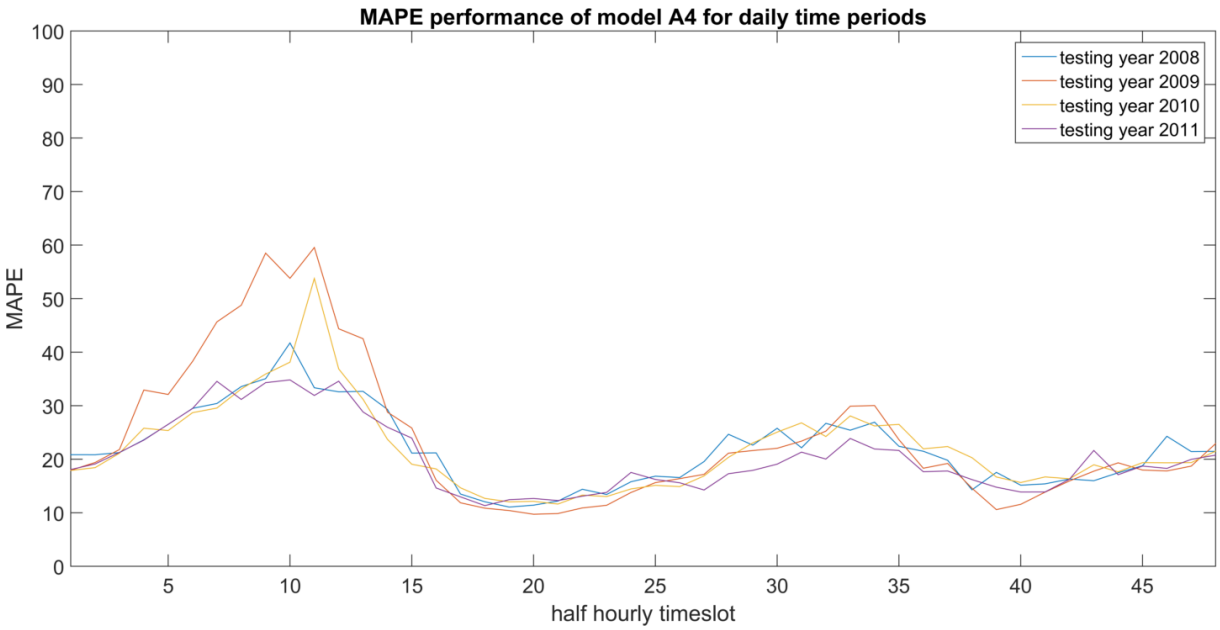

Figure 5-13: MAPE over each cross validation fold over each daily timeslot using model A4.

The accuracy of the total energy consumption in each fold is shown in Table 5-9. Once again the results here show very similar performance to that obtained by all other models in group A.

Table 5-9: Percentage error of model A4 in estimating yearly energy consumption within the training intervals.

Year	Total observed energy consumption [MWh]	Total predicted energy consumption [MWh]	Percentage error (%)	Over/under predicted
2008	670.8	650.27	3.06	Under
2009	739.14	718.06	2.85	Under
2010	727.01	747.71	-2.85	Over
2011	734.26	754.74	-2.79	Over

5.5 Group B case study

5.5.1 Model B1

5.5.1.1 Overview

Model set B1 consists of a single model targeting any day of the year, any day of the week and any time period of the day. The model accepts day of year, day of week and time of day as inputs. The entire data set is used for training and testing. Model B1 is similar to model A1, but unlike model A1 it does not take temperature as an input.

5.5.1.2 *Model outputs*

The output profile of model B1 for fold 4 of cross validation training is shown as a heat-map in Figure 5-14 and as a 3D mesh in Figure 5-15. The outputs for fold 1 to 3 are shown in Appendix A.5. Even without the temperature input, the model profile correctly exhibits many of the characteristics of the observed data, much the same as model A1. Model B1 shows less variance in the daily profile than model A1.

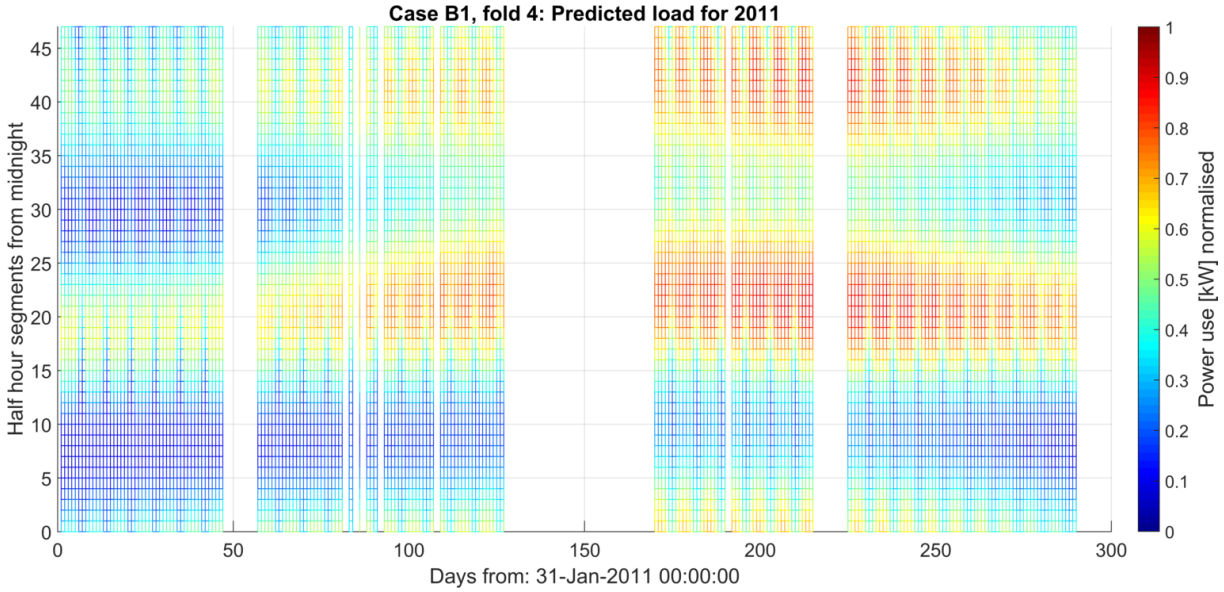

Figure 5-14: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model B1.

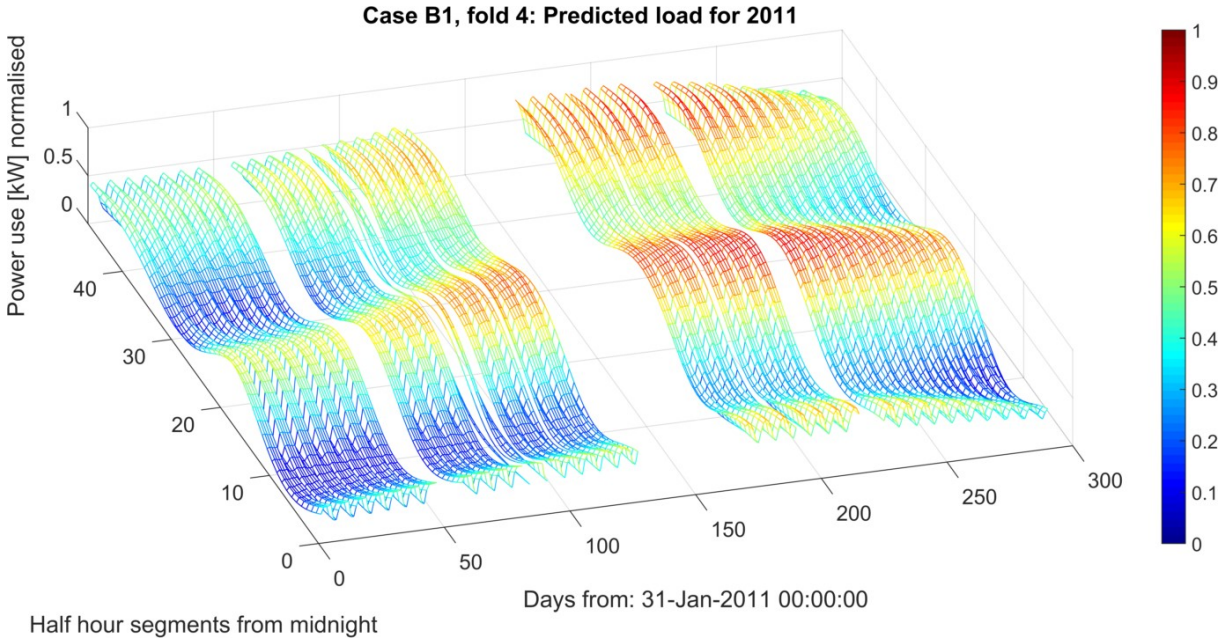

Figure 5-15: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model B1.

5.5.1.3 Comprehensive accuracy assessment

The accuracies of model B1 are shown in Table 5-10. All accuracies obtained are marginally better than found with model A1. This could be because model complexity is reduced by elimination of an input, or that the day to day behaviour of the temperature doesn't play such an important role in determining load behaviour, or both. Figure 5-16 shows the MAPE over each timeslot. Strangely, the performance over the early morning remains similar to that of model A1 for 2008 and 2001, but shows improved performance for 2009 and 2010.

Table 5-10: MAPE of each cross validation fold over each weekday category using model B1.

Fold	Testing year	MAPE (%)			
		Whole week	Weekday	Saturdays	Sundays
1	2008	24.84	24.37	26.59	25.46
2	2009	24.25	23.39	26.88	25.96
3	2010	22.14	21.29	24.55	24.18
4	2011	22.02	22.25	20.26	22.57
Mean		23.31	22.83	24.57	24.54

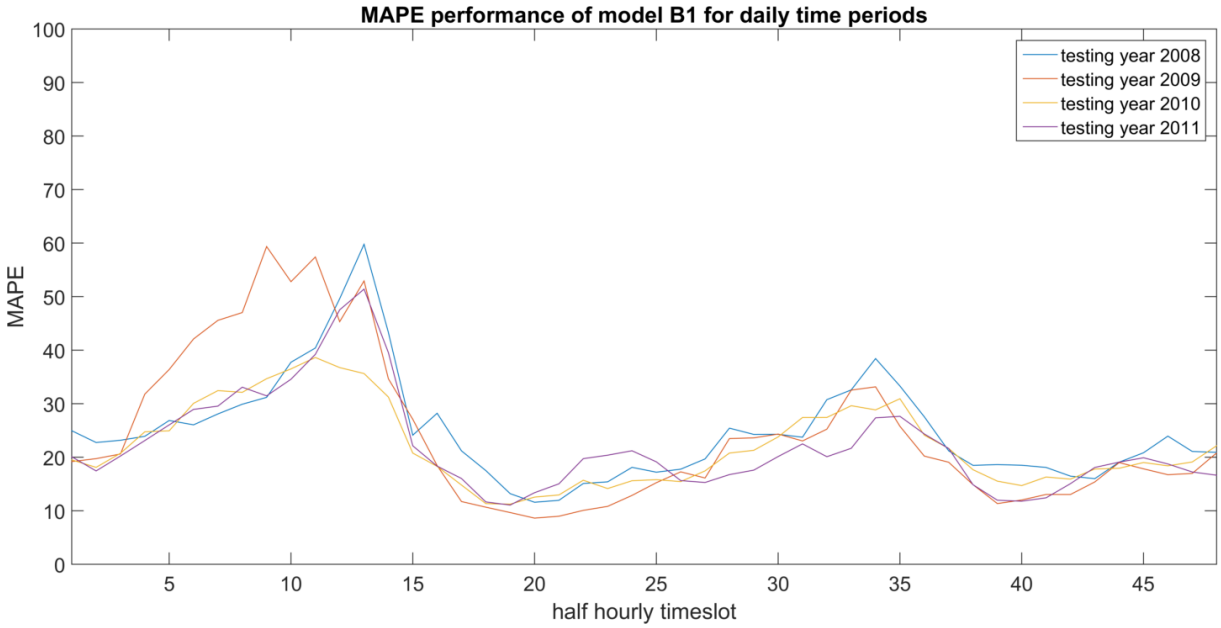

Figure 5-16: MAPE over each cross validation fold over each daily timeslot using model B1.

The accuracy of the total energy consumption in each fold is shown in Table 5-11. Compared to model A1 the percentage error for 2008 is worse, but better or the same for 2009 to 2011.

Table 5-11: Percentage error of model B1 in estimating yearly energy consumption within the training intervals.

Year	Total observed energy consumption [MWh]	Total predicted energy consumption [MWh]	Percentage error (%)	Over/under predicted
2008	670.8	646.44	3.63	Under
2009	739.14	726.01	1.78	Under
2010	727.01	746.7	-2.71	Over
2011	734.26	752.33	-2.46	Over

5.5.2 Model B2

5.5.2.1 Overview

Model set B2 consists of 3 models targeting weekdays, Saturdays and Sundays respectively. The models accept day of year and time of day as inputs. The data sets used for training and testing the models consist of subsets representing weekdays, Saturdays and Sundays respectively. Model B2 has a similar architecture to model A2, but without temperature as input.

5.5.2.2 Model outputs

The output profile of model B1 for fold 4 of cross validation training is shown as a heat-map in Figure 5-17 and as a 3D mesh in Figure 5-18. The outputs for fold 1 to 3 are shown in Appendix A.6. Similar to model A2 the daily profile of any given week does not differ much, though in model B2 to an even lesser extent.

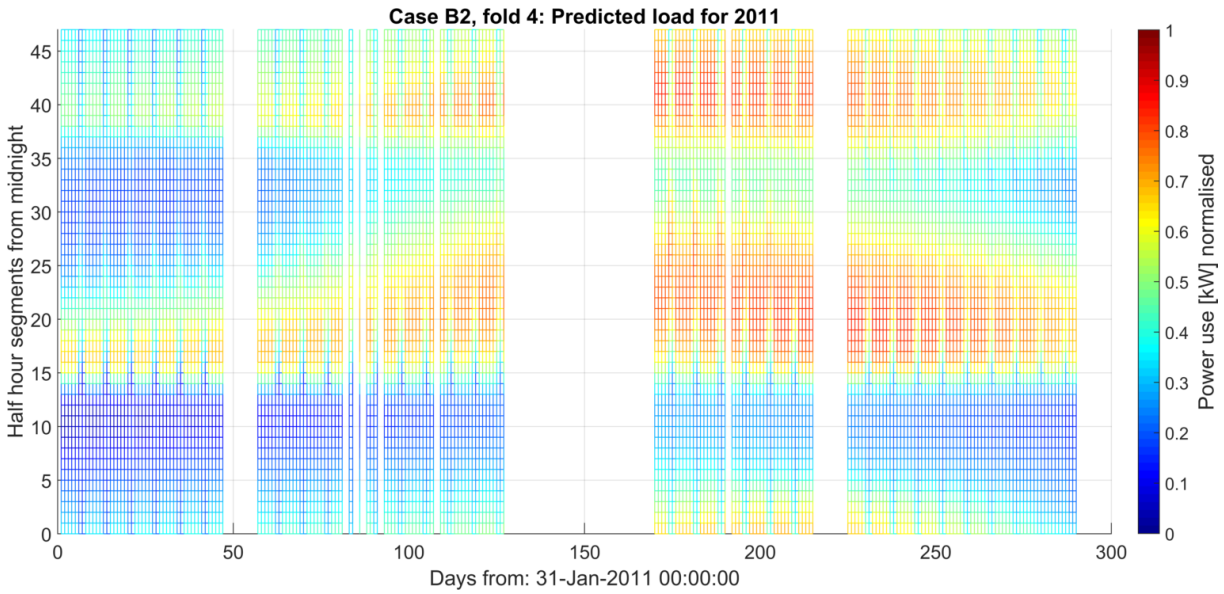

Figure 5-17: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model B2.

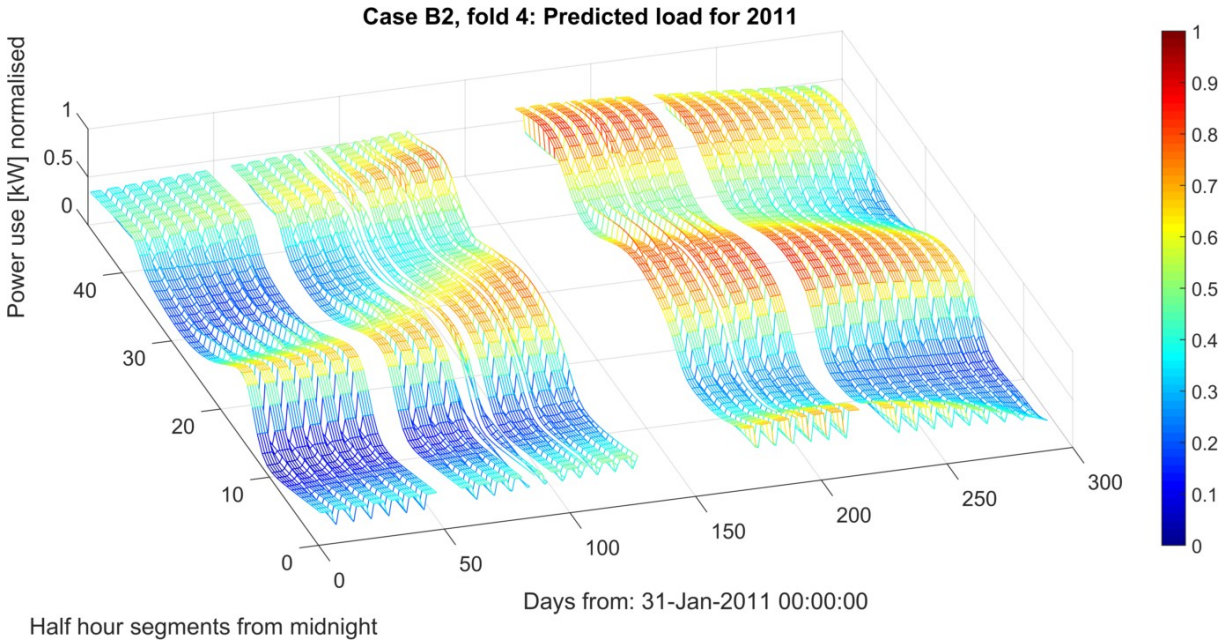

Figure 5-18: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model B2.

5.5.2.3 Comprehensive accuracy assessment

The accuracies of model B2 are shown in Table 5-12. The performance of this model is marginally better than that of model A2, with a total MAPE of 22.54 % as opposed to 24.19 %. So once again the model performance is increased by omission of the temperature input. The MAPE over each timeslot is visualised by Figure 5-19, which shows a sharp decrease in the prediction errors made between timeslot 10 and 15, when compared with that of models B1 and A2.

Table 5-12: MAPE of each cross validation fold over each weekday category using model B2.

Fold	Testing year	MAPE (%)			
		Whole week	Weekday	Saturdays	Sundays
1	2008	22.75	21.96	26.07	23.40
2	2009	24.91	24.05	26.54	27.59
3	2010	22.37	21.65	24.30	24.26
4	2011	20.13	19.78	19.86	22.27
Mean		22.54	21.86	24.19	24.38

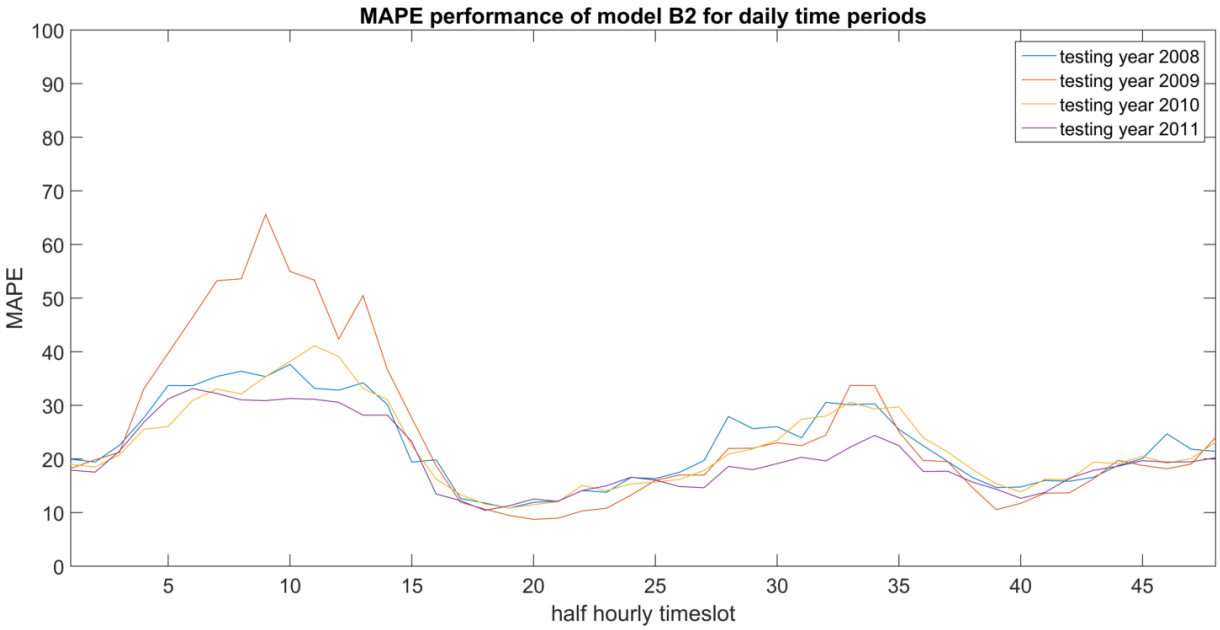

Figure 5-19: MAPE over each cross validation fold over each daily timeslot using model B2.

The accuracy of the total energy consumption in each fold is shown in Table 5-13. The performance in this regard is very similar to model B1.

Table 5-13: Percentage error of model B2 in estimating yearly energy consumption within the training intervals.

Year	Total observed energy consumption [MWh]	Total predicted energy consumption [MWh]	Percentage error (%)	Over/under predicted
2008	670.8	646.26	3.66	Under
2009	739.14	725.78	1.81	Under
2010	727.01	746.61	-2.7	Over
2011	734.26	752.55	-2.49	Over

5.5.3 Model B3

5.5.3.1 Overview

Model set B3 consists of 48 models targeting the 48 daily time periods respectively. The models accept day of the year and day of the week as inputs. The data sets used for training and testing the models consists of subsets representing the 48 daily time periods respectively. Model B3 has a similar architecture to model A3, but without temperature as input.

5.5.3.2 Model outputs

The output profile of model B3 for fold 4 of cross validation training is shown as a heat-map in Figure 5-20 and as a 3D mesh in Figure 5-21. The outputs for fold 1 to 3 are shown in Appendix A.7. Model B3 seems better suited than model B2 at following the high ramp rate of the load right before the morning peak, similar to the improvement seen in from model A2 to A3.

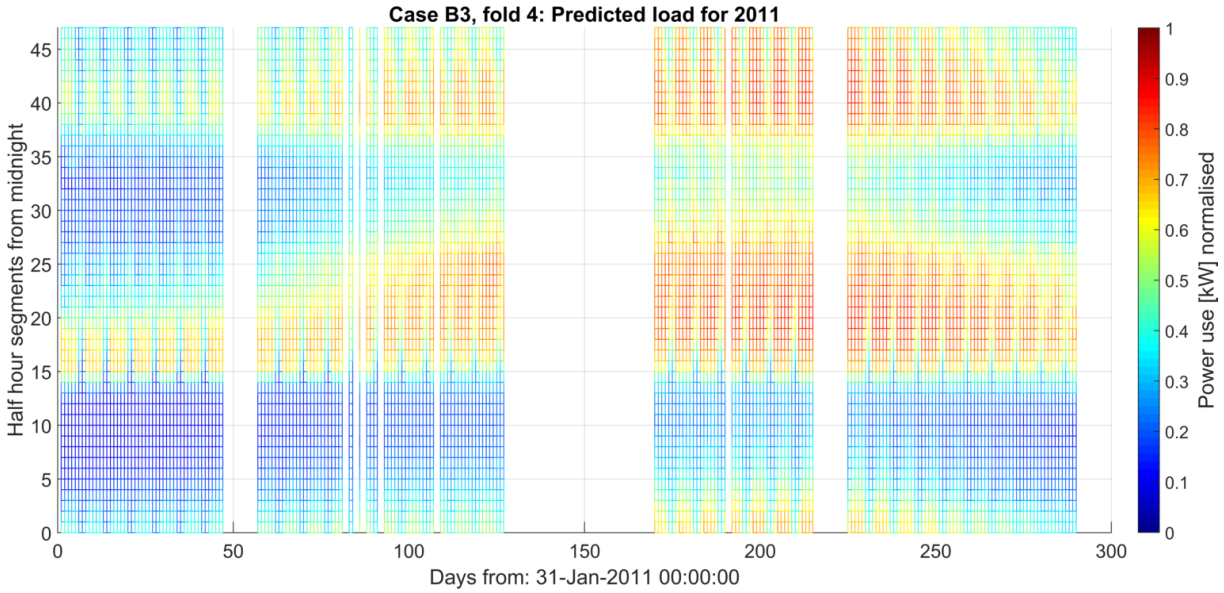

Figure 5-20: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model B3.

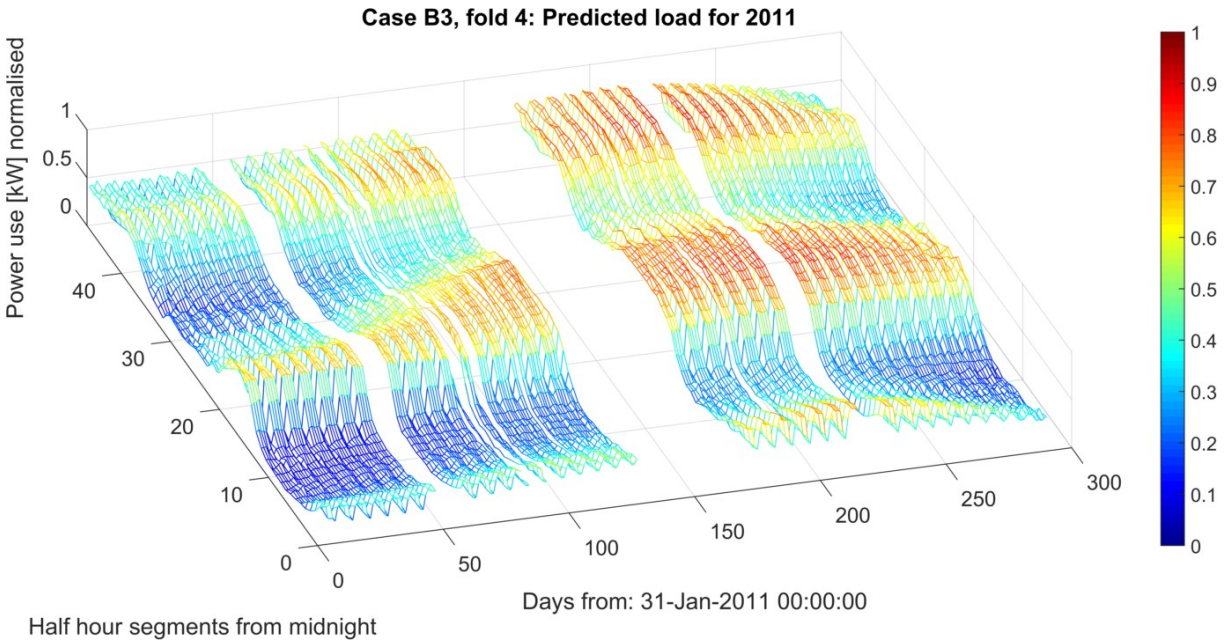

Figure 5-21: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model B3.

5.5.3.3 Comprehensive accuracy assessment

The accuracies of model B3 are shown in Table 5-14. Unlike models B1 and B2, this is the only model configuration in group B that does not outperform its group A counterpart. Instead, the accuracies are slightly worse for model B3 across the board when compared to A3, but still better than B1 and B2. The MAPE over each timeslot is visualised by Figure 5-22, and shows similar performance over all timeslots as model A3.

Table 5-14: MAPE of each cross validation fold over each weekday category using model B3.

Fold	Testing year	MAPE (%)			
		Whole week	Weekday	Saturdays	Sundays
1	2008	21.86	20.81	25.84	23.19
2	2009	23.40	22.39	26.44	25.46
3	2010	21.69	20.85	24.74	23.08
4	2011	19.49	18.87	20.37	21.90
Mean		21.61	20.73	24.35	23.41

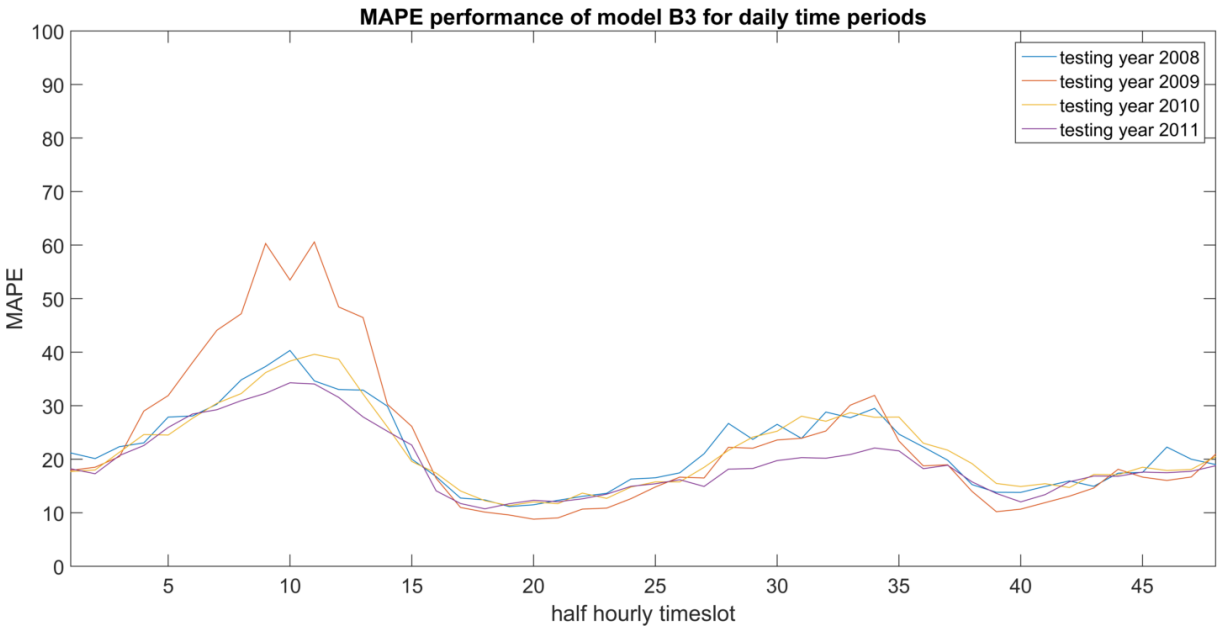

Figure 5-22: MAPE over each cross validation fold over each daily timeslot using model B3.

The accuracy of the total energy consumption in each fold is shown in Table 5-15, and is very similar to that obtained by the other models in group B.

Table 5-15: Percentage error of model B3 in estimating yearly energy consumption within the training intervals.

Year	Total observed energy consumption [MWh]	Total predicted energy consumption [MWh]	Percentage error (%)	Over/under predicted
2008	670.8	646.3	3.65	Under
2009	739.14	725.83	1.8	Under
2010	727.01	745.9	-2.6	Over
2011	734.26	752.41	-2.47	Over

5.6 Group C case study

5.6.1 Model C1

5.6.1.1 Overview

Model set C1 consists of a single model targeting any day of the year, any day of the week and any time period of the day. The model accepts temperature, day of week and time of day as inputs. The entire data set is used for training and testing. Model C1 has a similar architecture to model A1, but does not accept the day of year as input.

5.6.1.2 Model outputs

The output profile of model C1 for fold 4 of cross validation training is shown as a heat-map in Figure 5-23 and as a 3D mesh in Figure 5-24. The outputs for fold 1 to 3 are shown in Appendix A.8. The resemblance of the output profile to the observed data shows that prediction is possible with only the temperature input to serve as an indication of the time of the year. Like models A1 and B1, the daily profile is smooth and does a poor job of replicating the rapid ramp rate exhibited by the profile of the observed data. At a glance the model seems to under estimate the length of the morning peaks during for the colder periods of the year.

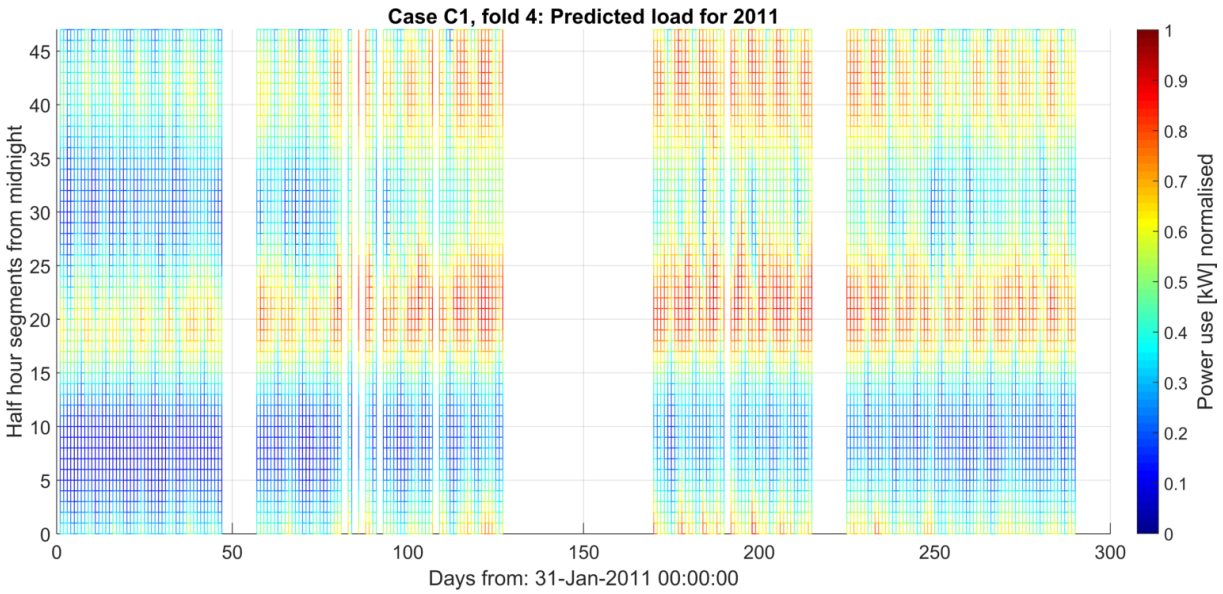

Figure 5-23: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model C1.

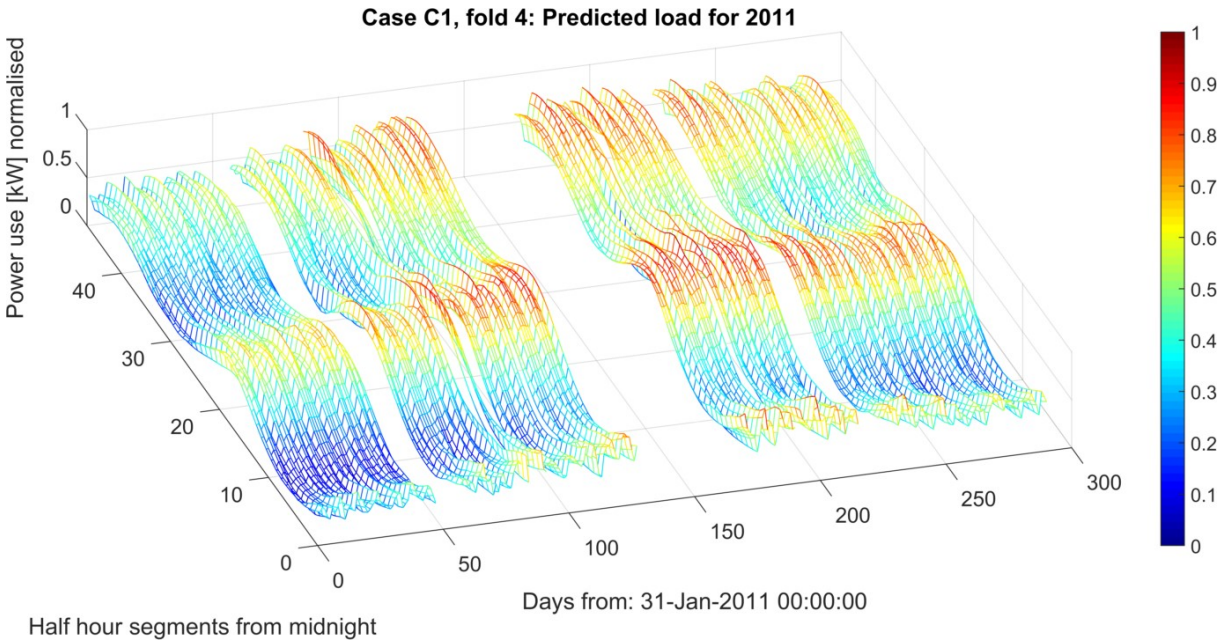

Figure 5-24: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model C1.

5.6.1.3 Comprehensive accuracy assessment

The accuracies of model C1 are shown in Table 5-16. With a mean MAPE of 26.61 % for all folds, model C1 has the poorest accuracy rating of this sub-model architecture, as models A1, B1, and D1 all perform better. The accuracies of the different weekday categories are more or less the same. The MAPE over each timeslot is visualised by Figure 5-25. Generally the model performs poorly from timeslot 3 to 16, which is consistent with models in group A and B. However this model exhibits a higher MAPE between timeslot 20 and 38 than is seen in group A and B.

Table 5-16: MAPE of each cross validation fold over each weekday category using model C1.

Fold	Testing year	MAPE (%)			
		Whole week	Weekday	Saturdays	Sundays
1	2008	26.56	26.71	25.25	27.10
2	2009	29.90	29.64	32.57	28.50
3	2010	25.06	24.73	26.68	25.16
4	2011	24.93	24.84	24.33	26.03
Mean		26.61	26.48	27.21	26.70

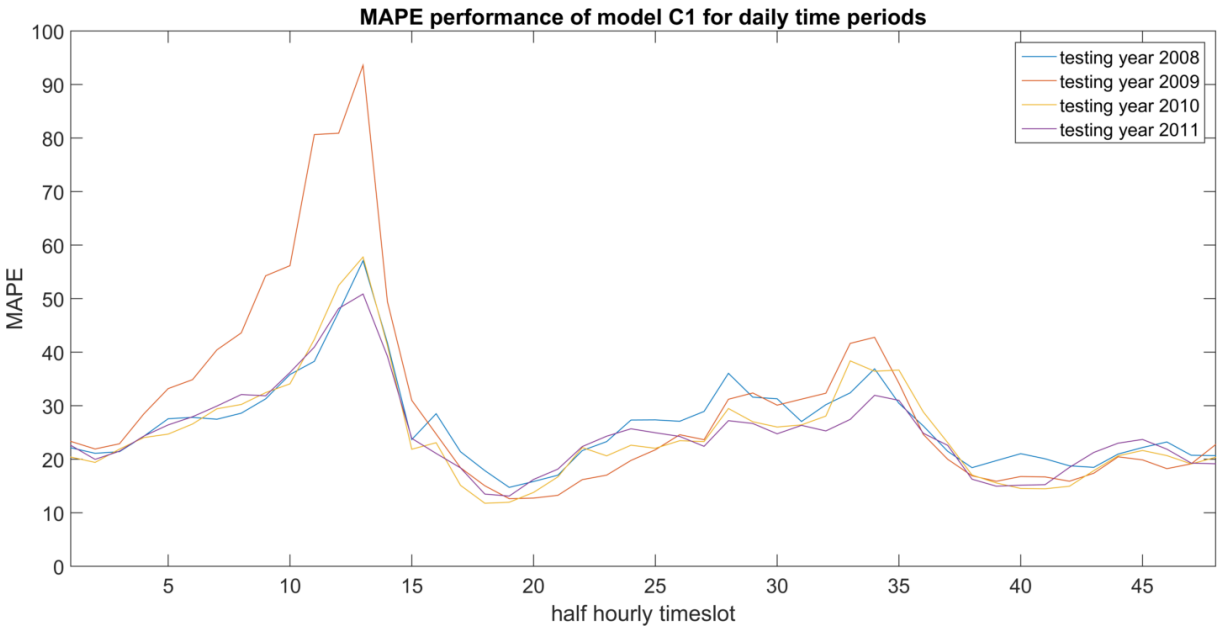

Figure 5-25: MAPE over each cross validation fold over each daily timeslot using model C1.

The accuracy of the total energy consumption in each fold is shown in Table 5-17. Just as all previously shown models, C1 under predicts the total energy for 2008 and 2009, and over predicts for 2010 and 2011. The percentage error for each year is the worst out of all other models so far.

Table 5-17: Percentage error of model C1 in estimating yearly energy consumption within the training intervals.

Year	Total observed energy consumption [MWh]	Total predicted energy consumption [MWh]	Percentage error (%)	Over/under predicted
2008	670.8	642.76	4.18	Under
2009	739.14	716.02	3.13	Under
2010	727.01	752.3	-3.48	Over
2011	734.26	760.52	-3.58	Over

5.6.2 Model C2

5.6.2.1 Overview

Model set C2 consists of 3 models targeting weekdays, Saturdays and Sundays respectively. The models accept temperature and time of day as inputs. The data sets used for training and testing the models consist of subsets representing weekdays, Saturdays and Sundays respectively. Model C2 has a similar architecture to model A2, but does not accept the day of year as input.

5.6.2.2 Model outputs

The output profile of model C2 for fold 4 of cross validation training is shown as a heat-map in Figure 5-26 and as a 3D mesh in Figure 5-27. The outputs for fold 1 to 3 are shown in Appendix A.9. The model output resembles that of model C1. Unlike models A2 and B2 there is slightly more distinction between the daily profiles of consecutive weekdays, though they are still less distinct than seen in model C1.

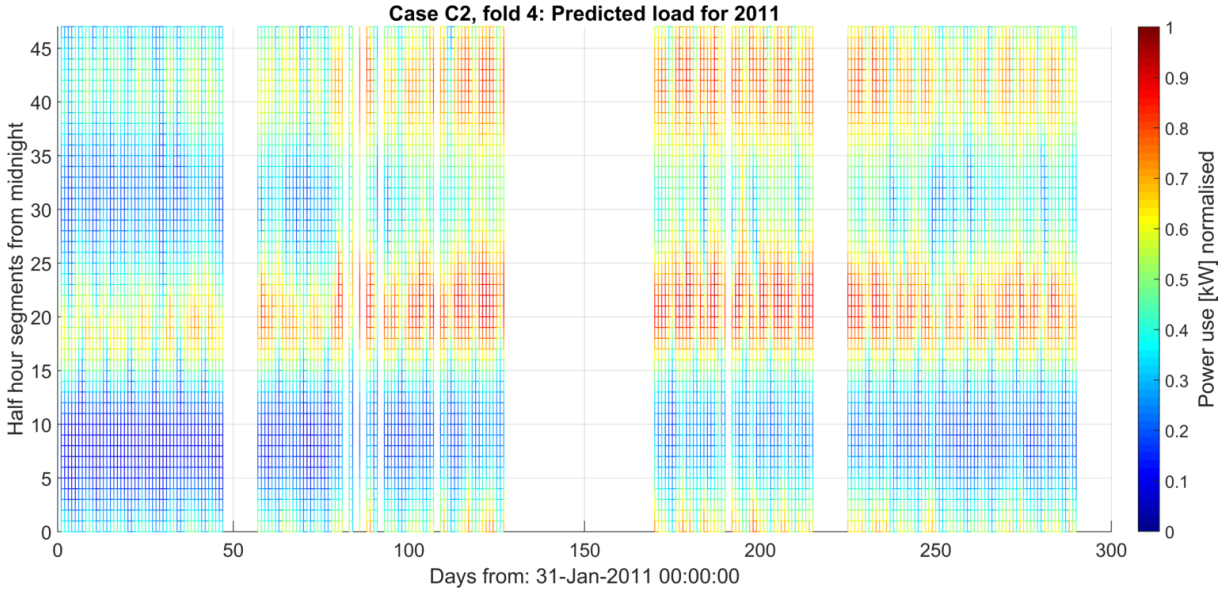

Figure 5-26: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model C2.

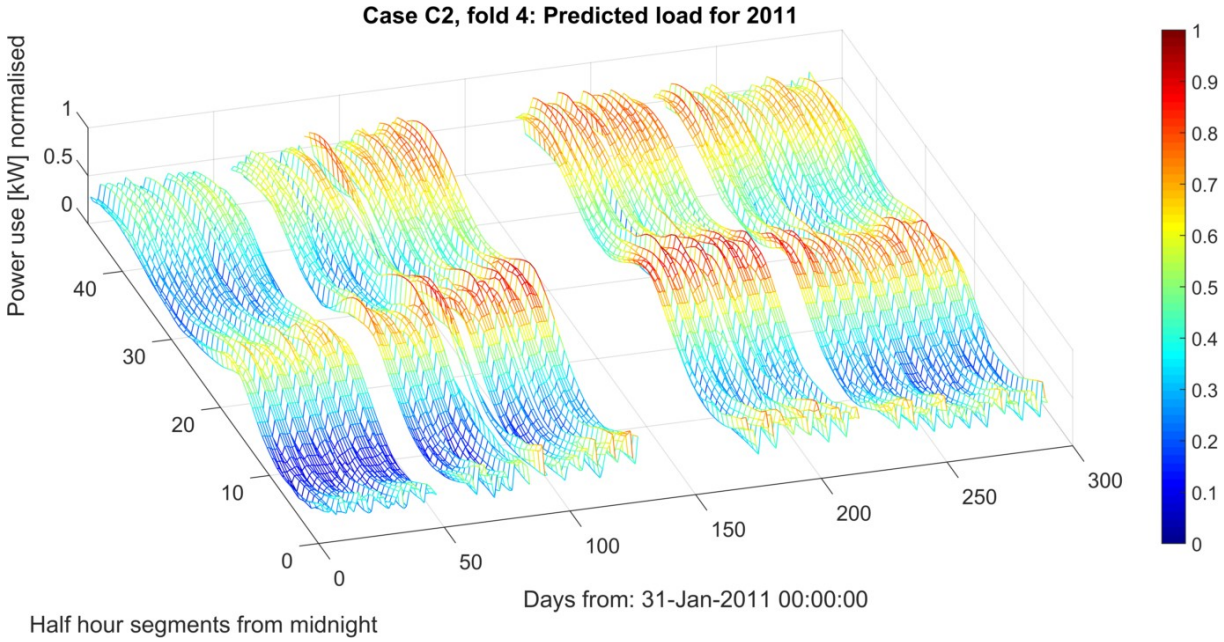

Figure 5-27: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model C2.

5.6.2.3 Comprehensive accuracy assessment

The accuracies of model C2 are shown in Table 5-18. The model seems to perform almost exactly the same as model C1, with no noticeable improvements to the MAPE error of any of the day of week categories. The MAPE over each timeslot is visualised by Figure 5-28, and once again the performance is the same as for model C1.

Table 5-18: MAPE of each cross validation fold over each weekday category using model C2.

Fold	Testing year	MAPE (%)			
		Whole week	Weekday	Saturdays	Sundays
1	2008	26.74	26.77	27.20	26.10
2	2009	29.25	29.30	30.17	28.09
3	2010	25.18	25.28	25.41	24.40
4	2011	24.82	25.23	22.89	24.57
Mean		26.50	26.65	26.42	25.79

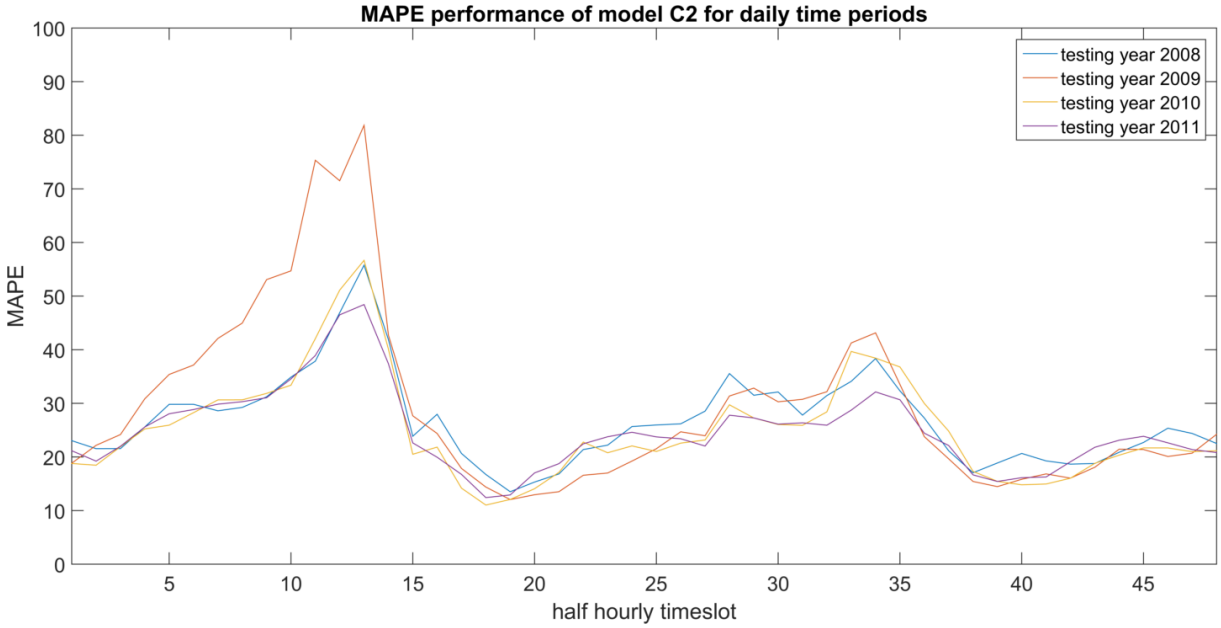

Figure 5-28: MAPE over each cross validation fold over each daily timeslot using model C2.

The accuracy of the total energy consumption in each fold is shown in Table 5-19, and shows yearly performance similar to model C1.

Table 5-19: Percentage error of model C1 in estimating yearly energy consumption within the training intervals.

Year	Total observed	Total predicted	Percentage error	Over/under

	energy consumption [MWh]	energy consumption [MWh]	(%)	predicted
2008	670.8	643.32	4.1	Under
2009	739.14	715.72	3.17	Under
2010	727.01	753.46	-3.64	Over
2011	734.26	760.01	-3.51	Over

5.6.3 Model C3

5.6.3.1 Overview

Model set C3 consists of 48 models targeting the 48 daily time periods respectively. The models accept temperature and day of the week as inputs. The data sets used for training and testing the models consists of subsets representing the 48 daily time periods respectively. Model C2 has a similar architecture to model A2, but does not accept the day of year as input.

5.6.3.2 Model outputs

The output profile of model C3 for fold 4 of cross validation training is shown as a heat-map in Figure 5-29 and as a 3D mesh in Figure 5-30. The outputs for fold 1 to 3 are shown in Appendix A.10. As with model A3 and B3, having a sub-model to target each timeslot individually results in an output profile which most closely resembles the observed data profile than the other models in group C.

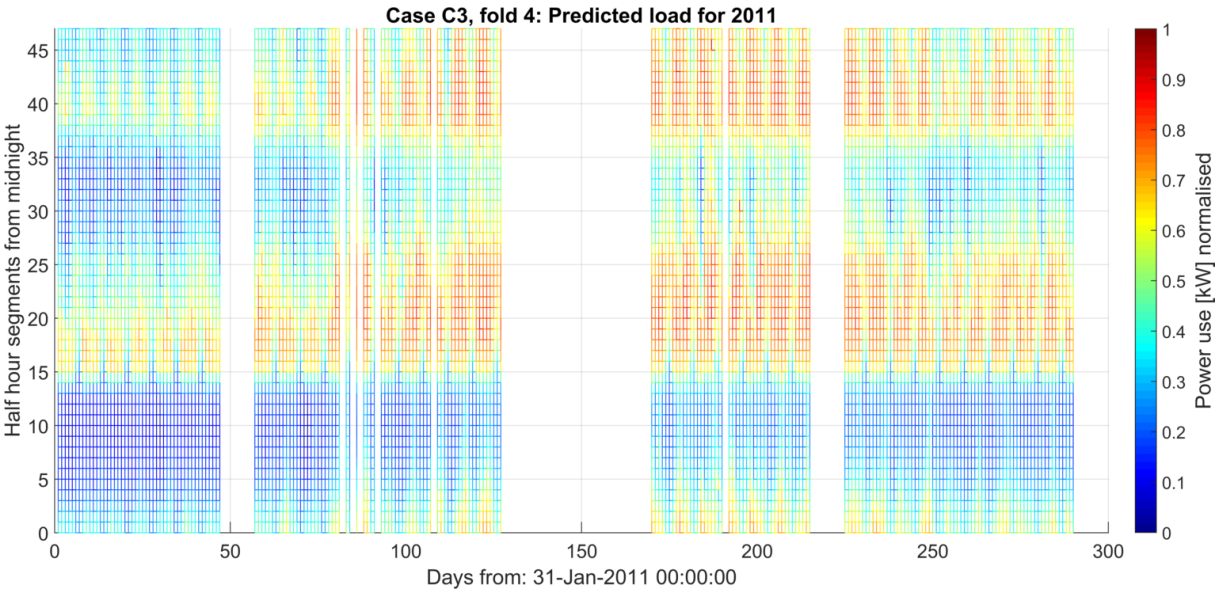

Figure 5-29: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model C3.

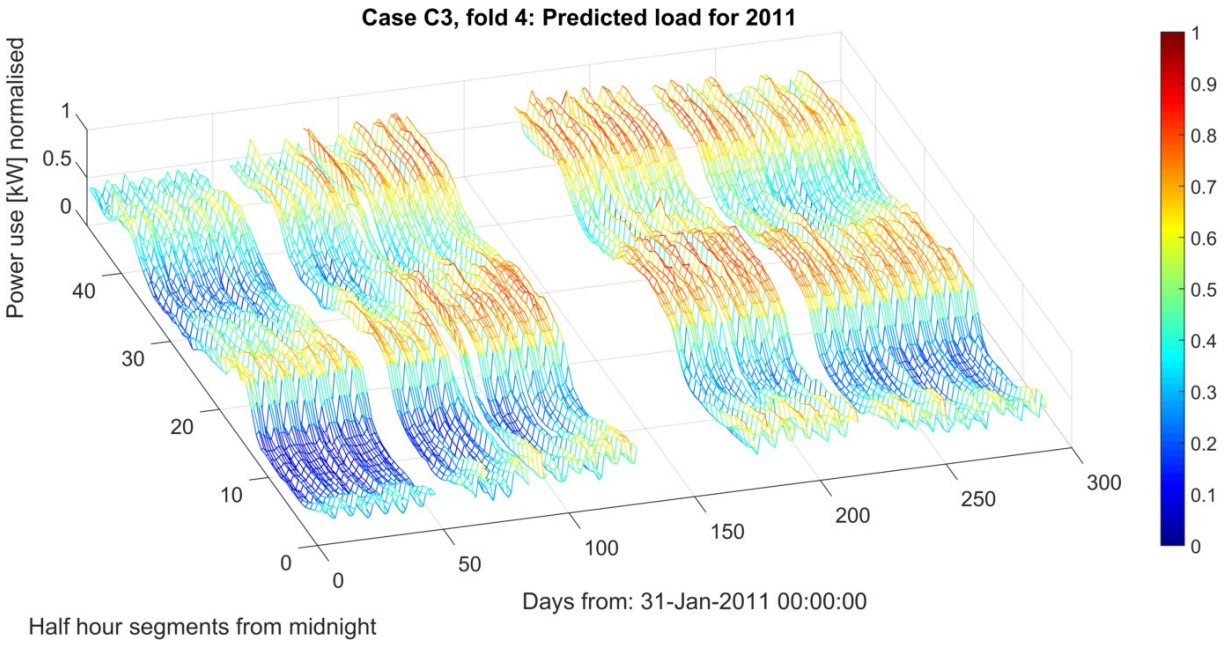

Figure 5-30: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model C3.

5.6.3.3 Comprehensive accuracy assessment

The accuracies of model C3 are shown in Table 5-20. The mean MAPE for all folds is 23.91 %, making it the best performing model in group C, and it can even contest models A1, A2, and B1. The prediction accuracy is higher for weekday periods, at a MAPE of 23.35 %, but lower for weekends by roughly 2 %. The MAPE over each timeslot is visualised by Figure 5-31.

Table 5-20: MAPE of each cross validation fold over each weekday category using model C3.

Fold	Testing year	MAPE (%)			
		Whole week	Weekday	Saturdays	Sundays
1	2008	24.13	23.56	25.14	25.99
2	2009	26.34	25.55	29.18	27.48
3	2010	22.65	22.06	24.99	23.38
4	2011	22.53	22.20	22.51	24.29
Mean		23.91	23.35	25.45	25.28

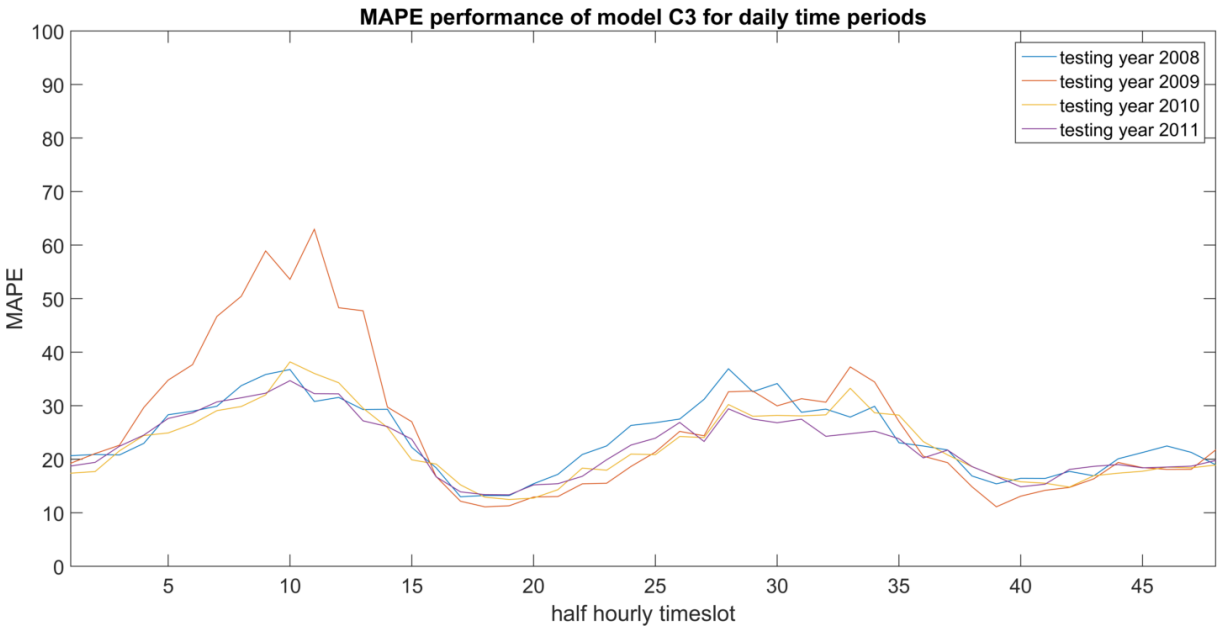

Figure 5-31: MAPE over each cross validation fold over each daily timeslot using model C3.

The accuracy of the total energy consumption in each fold is shown in Table 5-21, and is consistent with the performance of the rest of group C.

Table 5-21: Percentage error of model C3 in estimating yearly energy consumption within the training intervals.

Year	Total observed energy consumption [MWh]	Total predicted energy consumption [MWh]	Percentage error (%)	Over/under predicted
2008	670.8	641.93	4.3	Under
2009	739.14	716.07	3.12	Under
2010	727.01	753.13	-3.59	Over
2011	734.26	760.68	-3.6	Over

5.7 Group D case study

5.7.1 Model D1

5.7.1.1 Overview

Model set D1, like model set A1, consists of a single model targeting any day of the year, any day of the week and any time period of the day. The model accepts the same inputs as model set A1 (i.e. temperature, day of year, day of week and time of day), whereas the temperature input in model set A1 represents the ambient temperature at the given time of day, the

temperature input for model set D1 represents the ambient temperature 30 minutes prior to the given time of day. The entire data set is used for training and testing.

5.7.1.2 Model outputs

The output profile of model D1 for fold 4 of cross validation training is shown as a heat-map in Figure 5-32 and as a 3D mesh in Figure 5-33. The outputs for fold 1 to 3 are shown in Appendix A.11. Very little distinction can be made between the output profile of model D1 and A1.

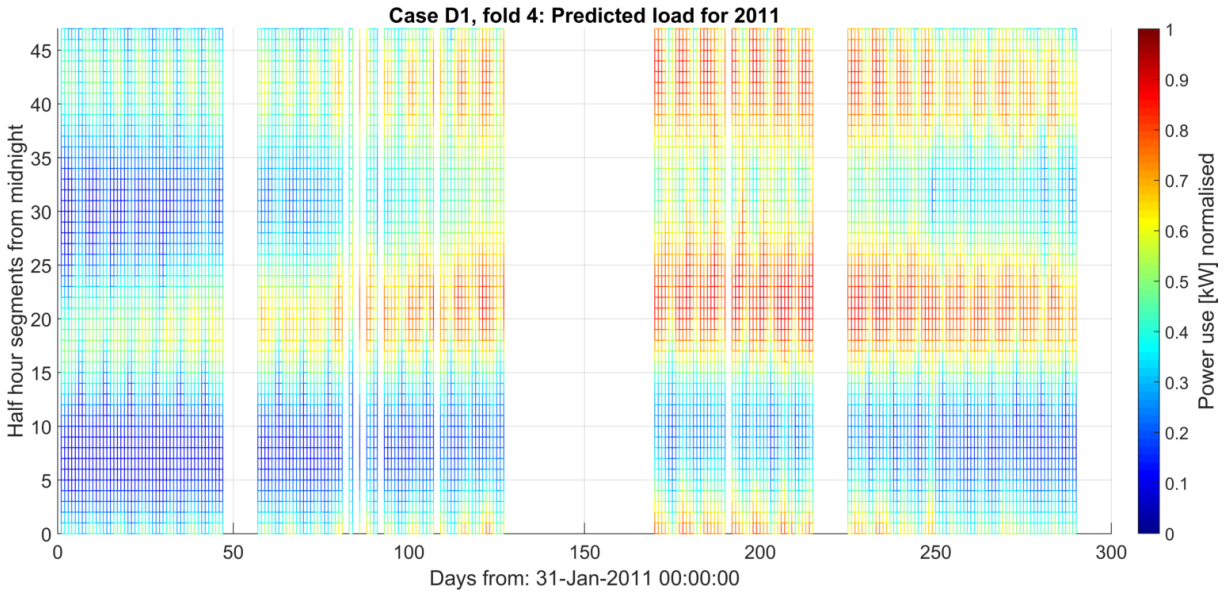

Figure 5-32: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model D1.

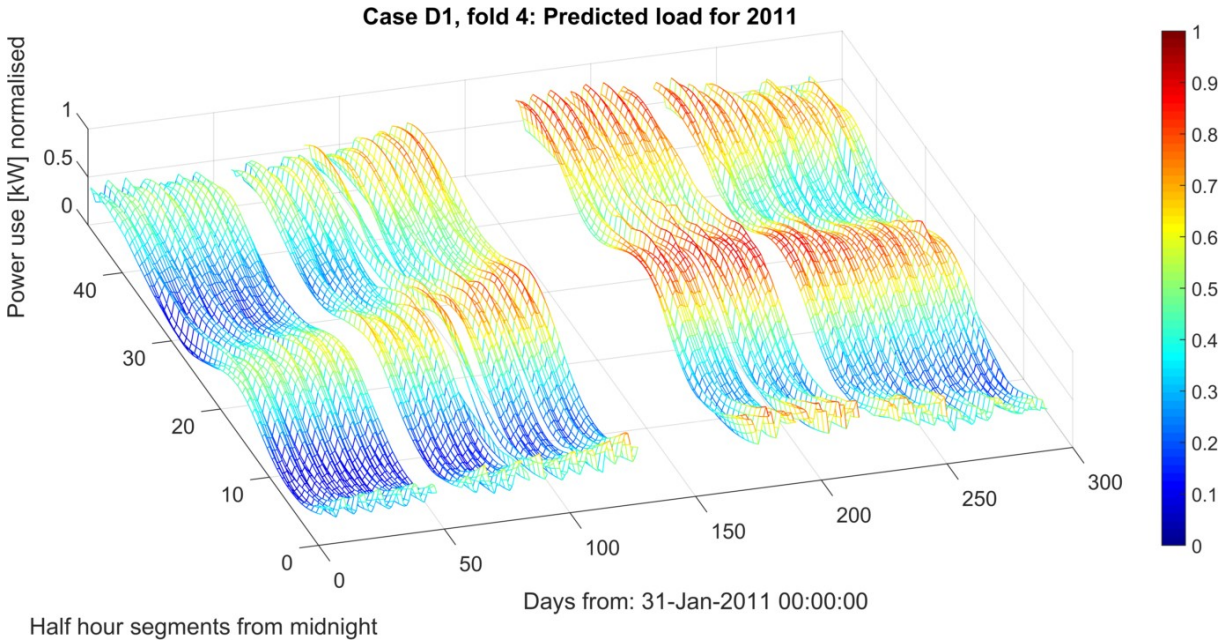

Figure 5-33: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model D1.

5.7.1.3 Comprehensive accuracy assessment

The accuracies of model D1 are shown in Table 5-22. The MAPE values are almost the same as model A1, except for slightly improved model performance for weekend period prediction. However this improvement to weekend performance by model D1 is very slight, and does not translate to a significant improvement in the mean MAPE for all folds. The MAPE over each timeslot, shown by Figure 5-34, is virtually identical to that of model A1.

Table 5-22: MAPE of each cross validation fold over each weekday category using model D1.

Fold	Testing year	MAPE (%)			
		Whole week	Weekday	Saturdays	Sundays
1	2008	24.11	23.56	25.89	25.12
2	2009	26.94	26.88	27.82	26.35
3	2010	24.99	25.03	25.76	24.01
4	2011	22.02	22.13	20.28	23.17
Mean		24.51	24.40	24.94	24.66

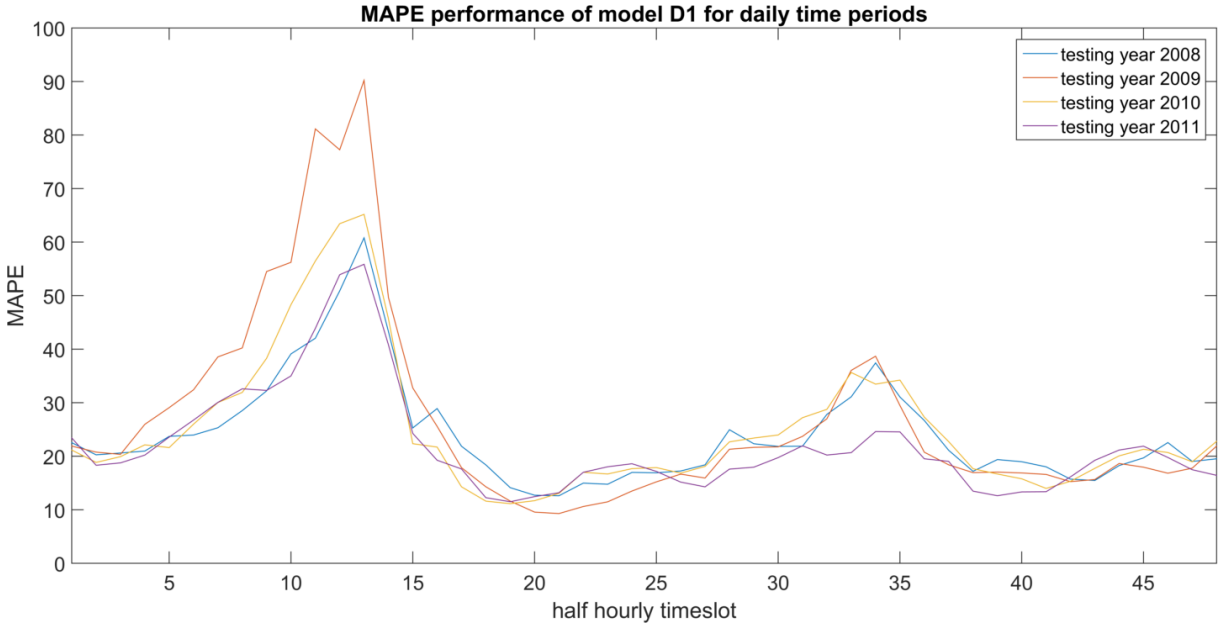

Figure 5-34: MAPE over each cross validation fold over each daily timeslot using model D1.

The accuracy of the total energy consumption in each fold is shown in Table 5-23.

Table 5-23: Percentage error of model D1 in estimating yearly energy consumption within the training intervals.

Year	Total observed	Total predicted	Percentage error	Over/under

	energy consumption [MWh]	energy consumption [MWh]	(%)	predicted
2008	670.8	649.31	3.2	Under
2009	739.14	718.01	2.86	Under
2010	727.01	747.38	-2.8	Over
2011	734.26	756.55	-3.04	Over

5.7.2 Model D2

5.7.2.1 Overview

Model set D2 consists of 3 models targeting weekdays, Saturdays and Sundays respectively. The models accept the same inputs as model set A2 (i.e. temperature, day of year, and time of day), whereas the temperature input in model set A2 represents the ambient temperature at the given time of day, the temperature input for model set D2 represents the ambient temperature 30 minutes prior to the given time of day. The data sets used for training and testing the models consist of subsets representing weekdays, Saturdays and Sundays respectively.

5.7.2.2 Model outputs

The output profile of model D2 for fold 4 of cross validation training is shown as a heat-map in Figure 5-35 and as a 3D mesh in Figure 5-36. The outputs for fold 1 to 3 are shown in Appendix A.12. At a glance the output profile generated by model D2 looks the same as that of model A2.

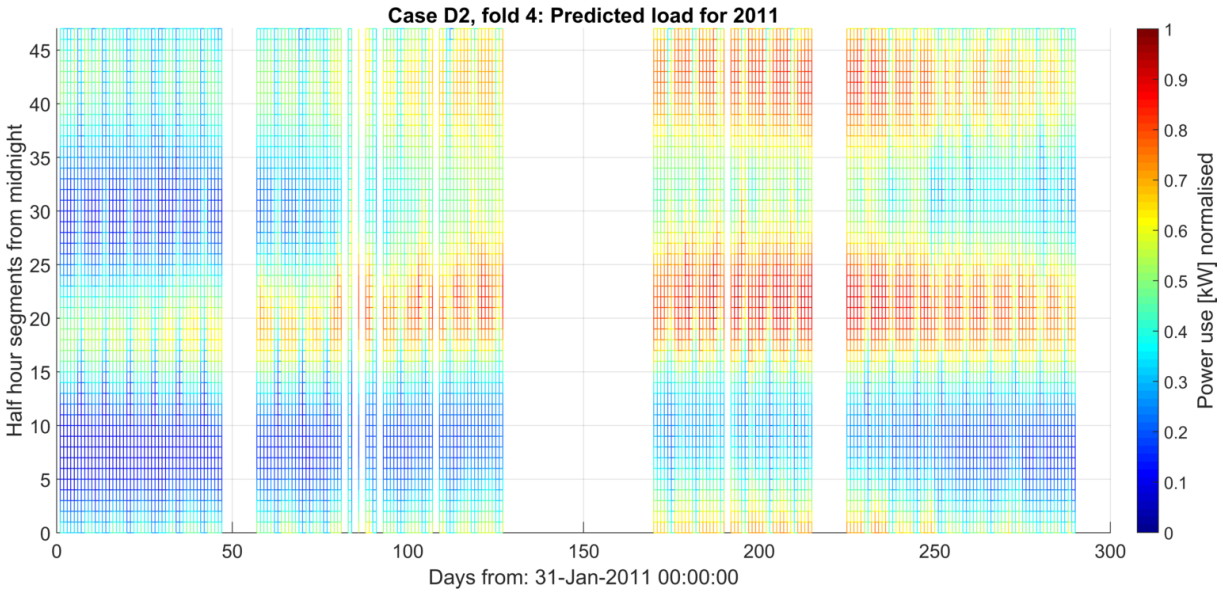

Figure 5-35: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model D2.

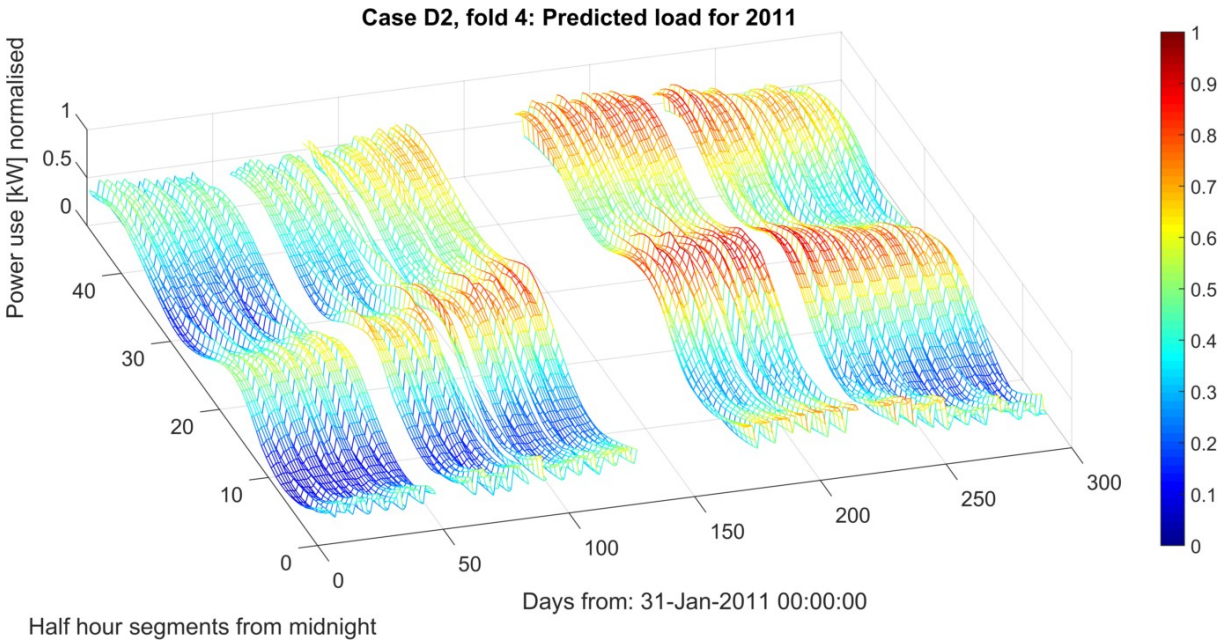

Figure 5-36: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model D2.

5.7.2.3 Comprehensive accuracy assessment

The accuracies of model D2 are shown in Table 5-24. The MAPE over each timeslot is visualised by Figure 5-37. As with model D1, adding a 30 minute delay to the temperature input doesn't translate to an improvement on model performance of the group A counterpart.

Table 5-24: MAPE of each cross validation fold over each weekday category using model D2.

Fold	Testing year	MAPE (%)			
		Whole week	Weekday	Saturdays	Sundays
1	2008	24.14	23.43	25.77	26.09
2	2009	27.13	26.79	28.62	27.36
3	2010	24.01	24.02	24.68	23.29
4	2011	22.04	22.21	20.64	22.52
Mean		24.33	24.11	24.93	24.81

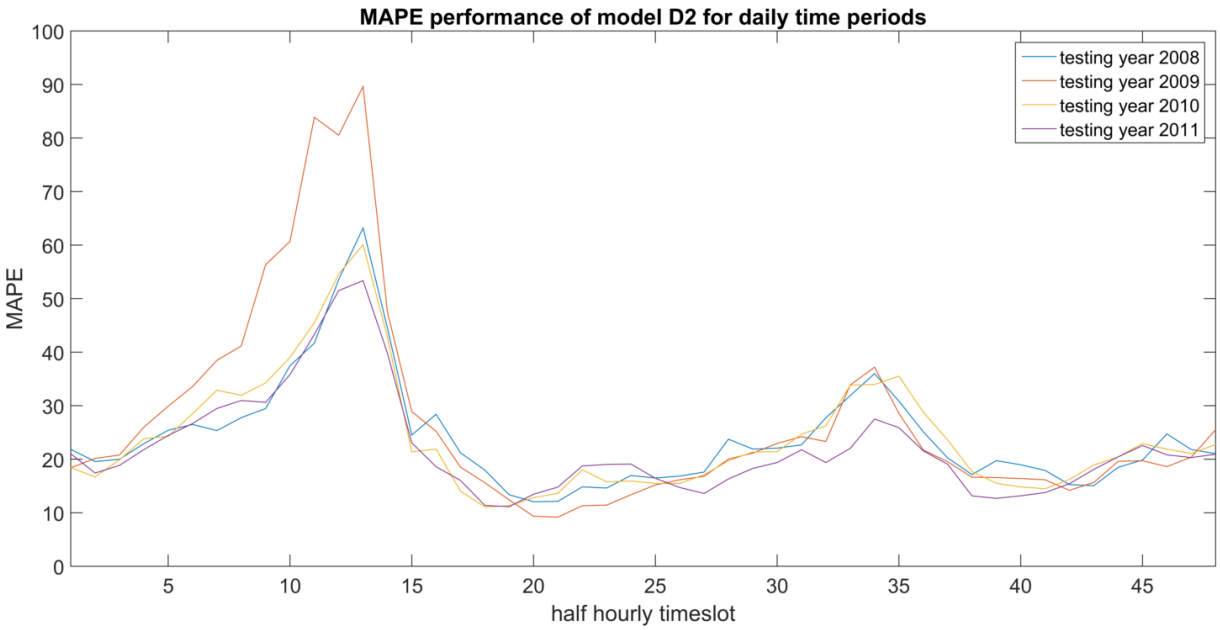

Figure 5-37: MAPE over each cross validation fold over each daily timeslot using model D2.

The accuracy of the total energy consumption in each fold is shown in Table 5-25.

Table 5-25: Percentage error of model D2 in estimating yearly energy consumption within the training intervals.

Year	Total observed energy consumption [MWh]	Total predicted energy consumption [MWh]	Percentage error (%)	Over/under predicted
2008	670.8	649	3.25	Under
2009	739.14	716.6	3.05	Under
2010	727.01	747.35	-2.8	Over
2011	734.26	755.58	-2.9	Over

5.7.3 Model D3

5.7.3.1 Overview

Model set D3 consists of 48 models targeting the 48 daily time periods respectively. The models accept temperature, day of the year and day of the week as inputs. The models accept the same inputs as model set A3 (i.e. temperature, day of year and day of week), whereas the temperature input in model set A3 represents the ambient temperature at the given time of day, the temperature input for model set D3 represents the ambient temperature 30 minutes prior to the given time of day. The data sets used for training and testing the models consists of subsets representing the 48 daily time periods respectively.

5.7.3.2 Model outputs

The output profile of model D3 for fold 4 of cross validation training is shown as a heat-map in Figure 5-38 and as a 3D mesh in Figure 5-39. The outputs for fold 1 to 3 are shown in Appendix A.13. As seems to be the case for all group D models, the profile is indistinguishable from that of its group A counterpart, model A3.

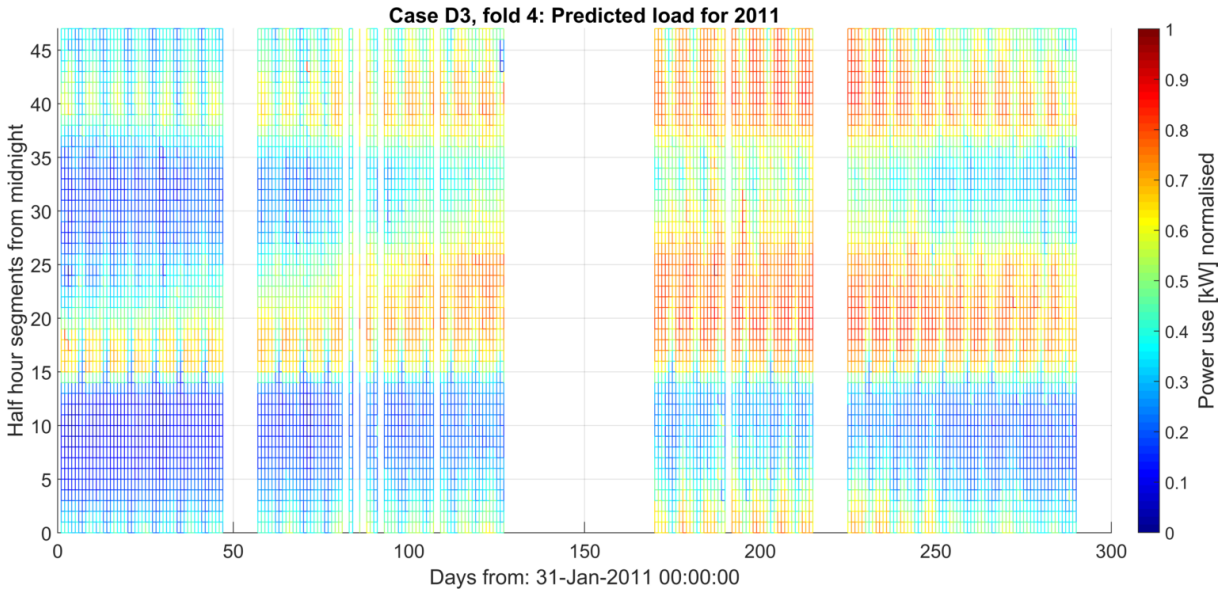

Figure 5-38: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model D3.

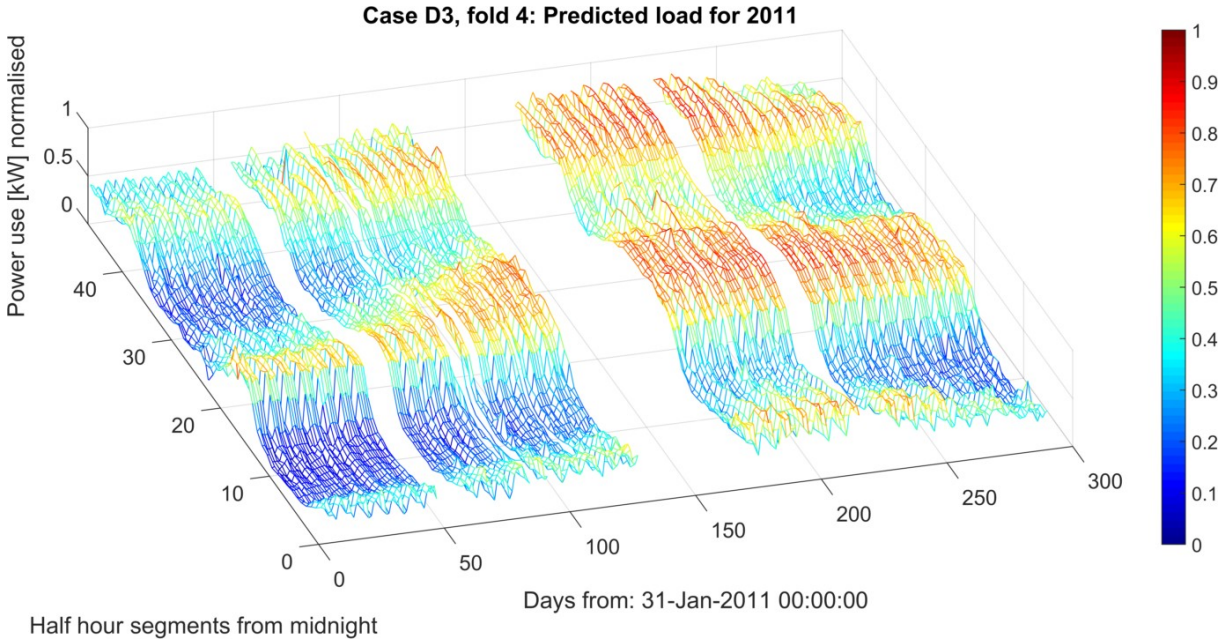

Figure 5-39: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model D3.

5.7.3.3 Comprehensive accuracy assessment

The accuracies of model D3 are shown in Table 5-26. The MAPE over each timeslot is visualised by Figure 5-40. As with group A, B and C, the third sub-model architecture results in the best model performance of each group. For the third time in group D it is shown that a delay in the temperature input has little effect on model performance.

Table 5-26: MAPE of each cross validation fold over each weekday category using model D3.

Fold	Testing year	MAPE (%)			
		Whole week	Weekday	Saturdays	Sundays
1	2008	20.87	19.71	24.62	23.01
2	2009	22.91	21.76	26.64	24.96
3	2010	20.92	19.87	24.82	22.55
4	2011	19.34	18.53	20.79	22.08
Mean		21.01	19.97	24.22	23.15

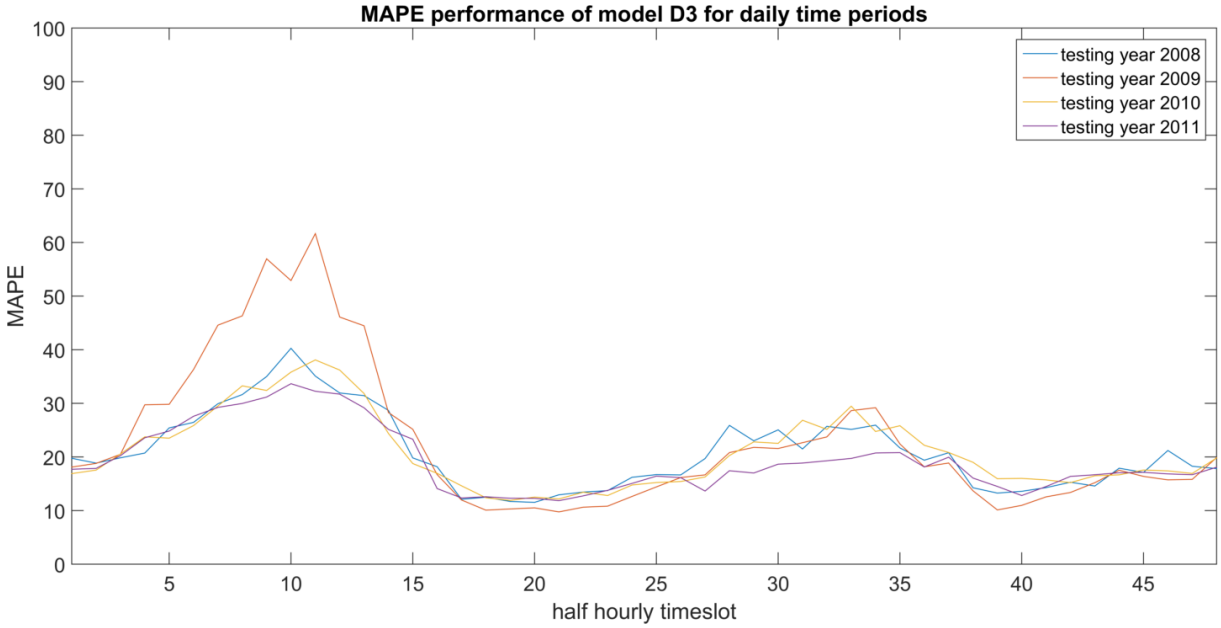

Figure 5-40: MAPE over each cross validation fold over each daily timeslot using model D3.

The accuracy of the total energy consumption in each fold is shown in Table 5-27.

Table 5-27: Percentage error of model D3 in estimating yearly energy consumption within the training intervals.

Year	Total observed energy consumption [MWh]	Total predicted energy consumption [MWh]	Percentage error (%)	Over/under predicted
2008	670.8	650.33	3.05	Under

2009	739.14	718.31	2.82	Under
2010	727.01	746.44	-2.67	Over
2011	734.26	756.03	-2.96	Over

5.7.4 Model D4

5.7.4.1 Overview

Model set D4 consists of a 3 subsets of models targeting weekdays, Saturdays and Sundays respectively. Each subset consists of 48 models targeting the 48 daily time periods respectively. The models accept temperature and day of the year as inputs. In this case the temperature input represents the ambient temperature 30 minutes prior to the time of day. The data sets used for training and testing the models consists of subsets representing the 3 day of week categories and associated 48 daily time periods respectively.

5.7.4.2 Model outputs

The output profile of model D4 for fold 4 of cross validation training is shown as a heat-map in Figure 5-41 and as a 3D mesh in Figure 5-42. This model has the same issue as model A4, as the sub-models that target weekend periods are trained with too few values, resulting in the random outlier predictions on weekends. The profile of model D4 can only be distinguished from that of model A4, because the outlier errors over weekends are different for each.

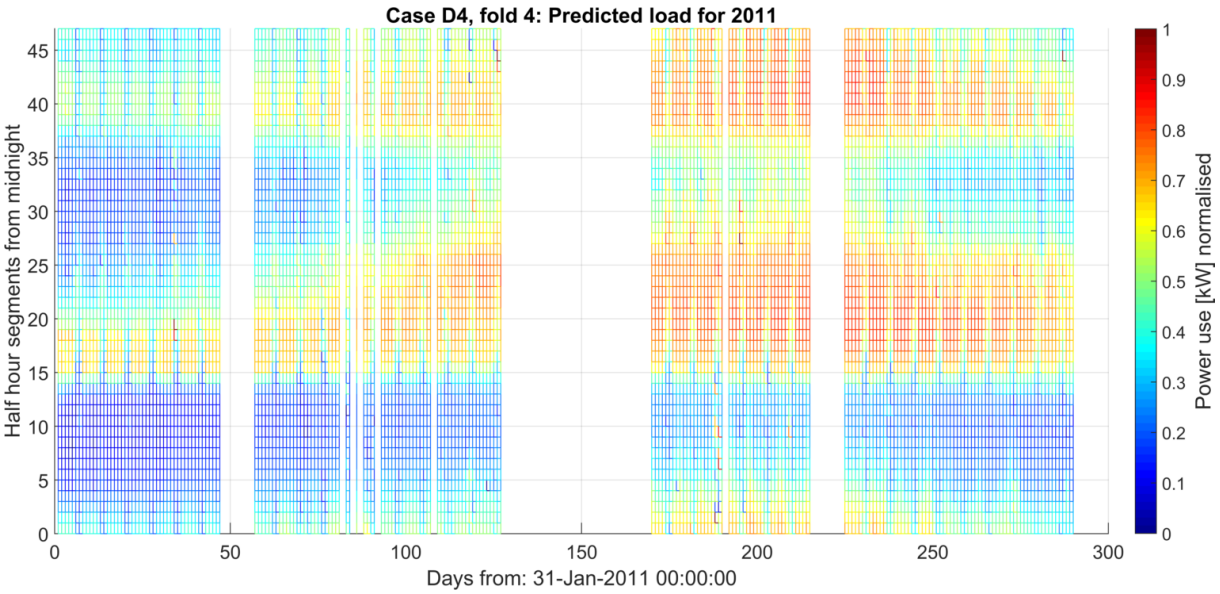

Figure 5-41: Heat map of predicted load profile for 2011, fold 4 of 4, as calculated by model D4.

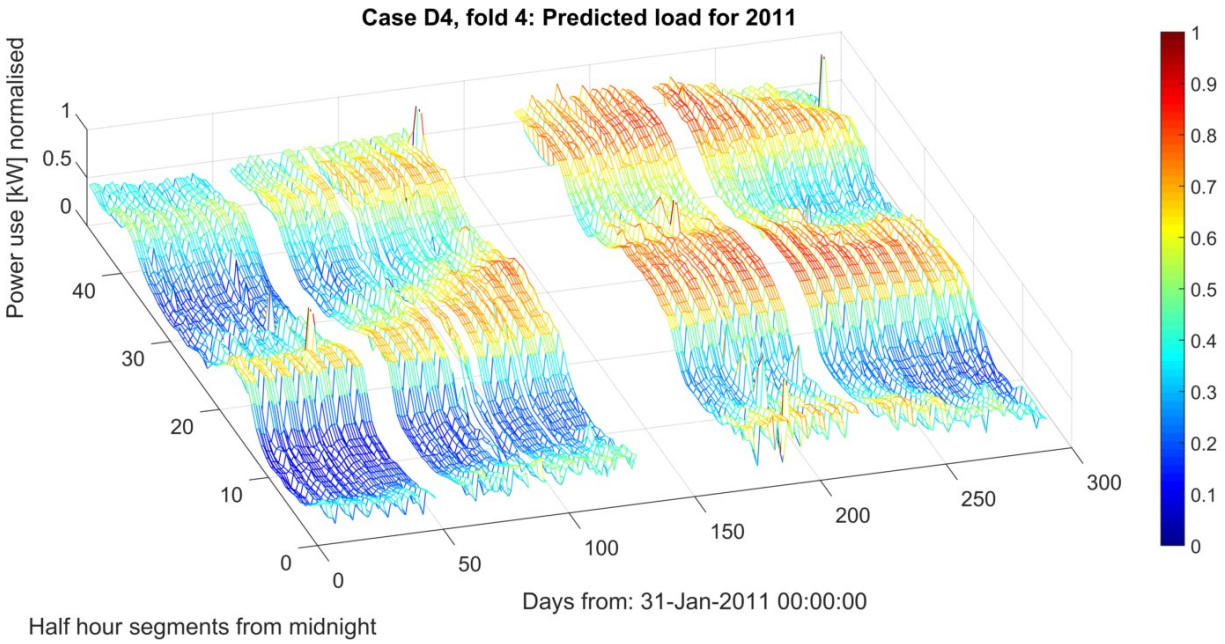

Figure 5-42: 3D mesh of predicted load profile for 2011, fold 4 of 4, as calculated by model D4.

5.7.4.3 Comprehensive accuracy assessment

The accuracies of model D4 are shown in Table 5-28. There is a slight increase over model A1 in accuracy for the total MAPE as well as for the MAPE of each day of the week category. The MAPE over each timeslot is visualised by Figure 5-43.

Table 5-28: MAPE of each cross validation fold over each weekday category using model D4.

Fold	Testing year	MAPE (%)			
		Whole week	Weekday	Saturdays	Sundays
1	2008	21.79	19.93	26.83	26.14
2	2009	23.48	21.87	27.89	27.25
3	2010	21.77	20.07	28.28	24.21
4	2011	20.02	18.70	23.40	23.55
Mean		21.77	20.14	26.60	25.29

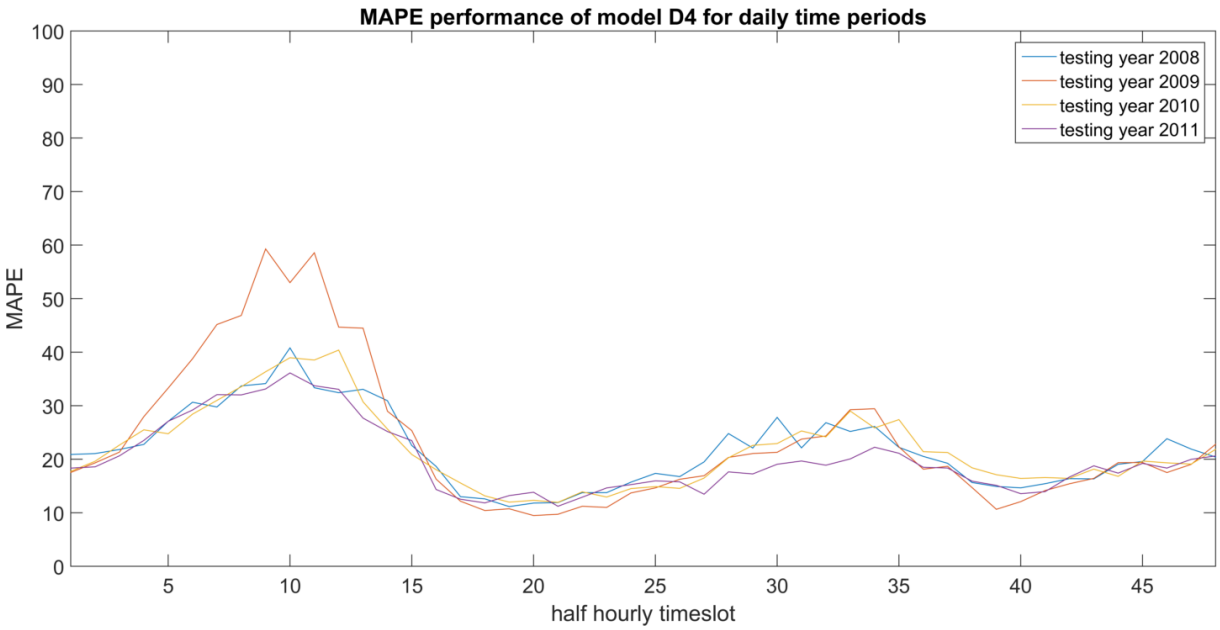

Figure 5-43: MAPE over each cross validation fold over each daily timeslot using model D4.

The accuracy of the total energy consumption in each fold is shown in Table 5-29.

Table 5-29: Percentage error of model D4 in estimating yearly energy consumption within the training intervals.

Year	Total observed energy consumption [MWh]	Total predicted energy consumption [MWh]	Percentage error (%)	Over/under predicted
2008	670.8	650.09	3.09	Under
2009	739.14	718.22	2.83	Under
2010	727.01	746.38	-2.66	Over
2011	734.26	755.58	-2.90	Over

5.8 Summary of case-study results

For every experimental load model Table 5-30 shows the average of the MAPE for each fold of cross validation verification. Also shown is the average MAPE of the distinct weekday categories. For Group A, B, C, and D, the third model configuration in each group shows the lowest MAPE. Load models A3 and D3 have practically the same MAPE of 21%, model B3 is only slightly worse at 21.61%, and Model C3 at 23.91%. Models A4 and D4 have low errors for weekday predictions, but high errors for weekend predictions.

Table 5-30: average MAPE of experimental load model cross validation.

Load model	Average k-fold MAPE (%)

	All	Weekday	Saturday	Sunday
A1	24.6	24.33	25.23	25.37
A2	24.19	23.85	25.35	24.79
A3	21.02	20.03	24.26	22.85
A4	21.96	20.21	26.83	26.08
B1	23.31	22.83	24.57	24.54
B2	22.54	21.86	24.19	24.38
B3	21.61	20.73	24.35	23.41
C1	26.61	26.48	27.21	26.7
C2	26.5	26.65	26.42	25.79
C3	23.91	23.35	25.45	25.28
D1	24.51	24.4	24.94	24.66
D2	24.33	24.11	24.93	24.81
D3	21.01	19.97	24.22	23.15
D4	21.77	20.14	26.6	25.29

5.9 Computational complexity

The computing complexity of a learning method generally goes up exponentially with an increase in the number of inputs [21]. For the case of ANFIS, refer to section 2.4.1 on its architecture. If the number of inputs goes up, the total number of input membership functions (MFs) go up, and the amount of possible rules goes up. With an increase in the number of rules there is an increase in the number of node function parameters that need to be updated during every epoch of training. Therefore an increase in the number of inputs leads to an exponential increase of mathematical operations performed during training. Because the *genfis2* routine is used to construct the initial FIS component of each individual ANFIS model, the number of MFs assigned to each input is automatically calculated and assigned based on the results of a subtractive clustering operation. Therefore the exact number of rules in each sub-model can vary, and a relative measure of computational complexity between the case models cannot be determined beforehand.

The fourteen load models presented in this chapter are trained using MATLAB, running on a machine with 16 GB of RAM, 2.4 GHz Intel i7 CPU, and a 256 MB SSD. Every individual ANFIS model or sub-model is trained over 800 epochs. Since every model is trained on the same hardware over the same amount of training iterations, the relative measure of the computational complexity of each case model can be determined by investigating the runtime of the training scripts.

The training runtime of each fold for each case model is shown in Table 5-31, along with the mean runtime of all four folds. Also shown in Table 5-31 is the number of distinct ANFIS routines employed in each case model, as well as the number of inputs in each of those sub-models. From this it is abundantly clear that the runtime of a case model depends primarily on the number of inputs per sub-model. The mean training runtime for case models is roughly 50 minutes with 4 inputs, 10 minutes with 3 inputs, and 2 to 5 minutes with 2 inputs. The number of distinct ANFIS structures that a case model is made up of has little effect, as an increase in the number of sub-models simply leads to proportional shrinking of the size of the training dataset of each sub-model. Therefore model A2, which is divided into 3 sub-models, is trained in the same time as model A3, which is divided into 48.

Table 5-31: Duration of training of each cross validation fold for each model.

Model:	**Number of distinct models**	**Inputs per sub-model**	**Model training runtime [minutes]**				
			Fold 1	**Fold 2**	**Fold 3**	**Fold 4**	**Mean**
A1	1	4	59.5	54.03	50.08	50.27	53.47
A2	3	3	10.37	10.52	10.27	11.13	10.57
A3	48	3	10.88	10.22	10.02	10.22	10.34
A4	144	2	2.38	2.15	2.25	2.17	2.24
B1	1	3	28.59	31.62	29.18	29.3	29.67
B2	3	2	5.75	5.65	5.5	5.73	5.66
B3	48	2	5.58	5	5.25	5.17	5.25
C1	1	3	12.78	13.9	12.17	12.2	12.76
C2	3	2	2.97	2.9	2.87	2.9	2.91
C3	48	2	3.43	3.37	3.37	3.27	3.36
D1	1	4	56.13	45.8	48.95	45.7	49.15
D2	3	3	9.58	9.72	8.88	9.9	9.41
D3	48	3	10.43	9.93	9.75	10.13	10.06
D4	144	2	2.22	2.1	2.15	2.17	2.16

5.10 Analysis of time-series prediction accuracies

5.10.1 Overview

In this section the behaviour of the load model accuracies are investigated through an analysis of the difference in the relative error size within the range of timeslots, as well as the range of weekday categories.

5.10.2 Accuracies of timeslots

There is a consistent behavioural trend between the timeslot accuracies of each load model in the case study. Consider Figure 5-4, wherein the MAPE obtained by model A1 in each timeslot is graphed for all folds. Notice that this MAPE profile has a high peak between timeslots 3 and 16 in the early morning, and a low peak between timeslots 30 and 37 in the evening. It is shown that the trend of this timeslot error profile holds true for every other case model. It is interesting to note that these two periods of poor timeslot accuracy coincide almost perfectly with the periods of the day generally associated with low or virtually no power use, i.e. the early morning before the morning peak, and the period between the morning and evening usage peaks. The timeslot accuracies associated with the periods of peak power use are generally much better.

It was found that this behaviour is not so much a result of exceptionally poor model performance over those intervals, but rather an unfortunate shortcoming of the MAPE accuracy measure when applied to predictions of very low values. For example if the observed load at a point is 0.01 and a model predicts a value of 0.04, what would seem like a very slight relative error translates to a percentage error of 400%. This effect is visualised in Figure 5-44, which shows the percentage errors of every data point predicted by model A1 over the size of the observed data for that point. The percentage error scatter plots of the other load models can be found in Appendix B.1. It is apparent that the percentage error is generally high for smaller values of the observed load. As the MAPE of each timeslot is calculated using data from the whole year, the first error peak is higher than the second, because the load is consistently small during that period throughout the year. The second timeslot MAPE peak is lower because this near-zero load between timeslots 30 and 37 only really occurs for the first semester of each year.

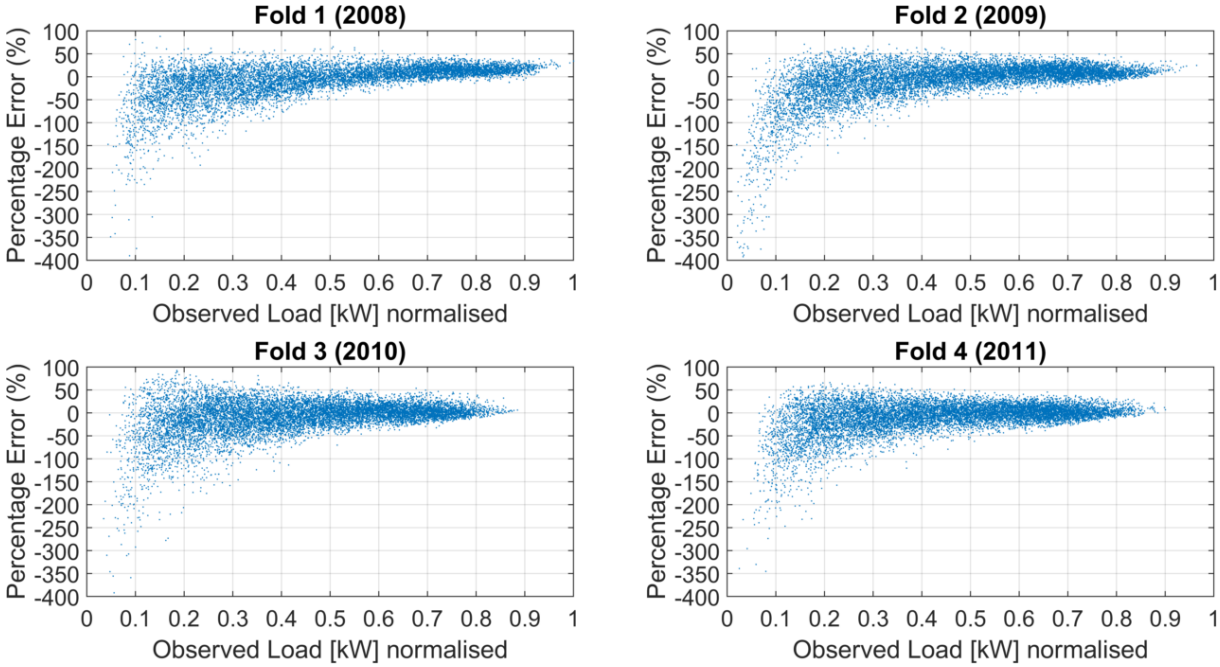

Figure 5-44: Percentage error of model A1 predictions over size of observed load at error.

5.10.3 Accuracies of weekday categories

Another observation that is consistent over all load models in the case study is that the accuracies of predictions of weekday load profiles are always better than those of Saturdays and Sundays. Weekdays make up 70% of the total data, and Saturdays and Sundays make up 15% respectively. Therefore it makes sense that the load models are better suited at predicting weekday profiles, as that is what the majority of the training data is made up of. It is interesting to note that this trend would probably persist to some degree for cases where there are many more years of data available for training, as the ratio of weekday to weekend data would remain skewed. It is therefore an inherent problem of load prediction that is dependent on the weekly cycle.

5.11 Analysis of total energy prediction accuracies

Each load model predicts a power use profile for every year of cross validation training. This profile, depicting the average power usage per half-hour, is used to calculate the energy consumption per half-hour. From this the total energy consumption of the year is calculated and can be compared with that of the actual load over the same testing intervals. The percentage error of this total energy prediction was shown for each load model case earlier in this chapter. The mean of the percentage error of the total energy prediction for each model is shown per year in Table 5-32, along with the standard deviation of these errors. The positive

mean of the percentage error of 3.45% for 2008, in combination with the relatively small standard deviation of 0.47% shows that all load models in the case study under-estimate the energy use in 2008. The same can be said of 2009, while the energy consumption in 2010 and 2011 is consistently over-estimated.

It is believed that this consistent under-, and over-estimation of specific years by all load models can be attributed to the yearly decrease in the average energy consumption from 2008 to 2010, shown in Table 3-5. For example when predicting the yearly profile for 2008 a load model is trained with data from 2009 to 2011, but because the average power consumption is higher for 2008 than any of the other years, the model under-estimates. The average power consumption of 2009 is higher than 2010 and 2011, but lower than 2008. Therefore predicting 2009 involves training with two years of data that exhibit lower values than the target year, and one that is higher. The model still under-estimates, but to a lesser degree than with 2008, which is trained with three years of data that exhibits lower values than the target year. The average power consumption is similar for 2010 and 2011, and as such the load models over-estimate the energy of both those years to an equal extent. All this is compelling evidence that most of the proposed load models are actually quite good at predicting the total energy consumption of specific periods, and that the percentage errors obtained are not so much errors as a percentage value of how much the nature of the target load differs from the training data. Although a longer range of consistent data would be needed to better support this conclusion.

Table 5-32: Mean and standard deviations of the percentage error for predictions of total energy use in each year, shown for each individual group as well as that of all models.

Prediction year	**Percentage error (%)**									
	Group A		**Group B**		**Group C**		**Group D**		**All**	
	mean	**STD**	**mean**	**STD**	**mean**	**STD**	**mean**	**STD**	**mean**	**STD**
2008	3.05	0.1	3.65	0.02	4.19	0.1	3.15	0.09	3.45	0.47
2009	2.87	0.04	1.8	0.02	3.14	0.03	2.89	0.11	2.70	0.51
2010	-2.76	0.06	-2.67	0.06	-3.57	0.08	-2.73	0.08	-2.91	0.37
2011	-2.84	0.06	-2.47	0.02	-3.56	0.05	-2.95	0.07	-2.95	0.38

Table 5-32 also shows the mean of the percentage error of the total energy prediction for each case study group. The standard deviation of the errors within the same group is never higher than 0.1%. This means that the reduction of model inputs along with the implementation of different sub-model architectures has little effect on the accuracy of the yearly energy

consumption. It is difficult to judge which of the load models in group A, B and D are better suited for yearly energy prediction, but it is clear that group C is the obvious loser in this regard.

Chapter 6: Conclusions and Recommendations

6.1 Conclusions

6.1.1 Overview

The conclusions of the investigation will be presented with reference to the original project objectives. These are as follows:

- Determine the feasibility of the application of ANFIS to model and predict the hot water energy consumption profiles associated with student residences for medium term forecasting horizons.
- Perform a comparative study in which different implementations of the ANFIS function are utilised with the objective of gaining insight into the relative strengths and weaknesses of the various implementations. This objective entails the following:
  - Determine whether partitioning the load model into multiple models that each operates on a subset of the given data improves accuracies and reduces the computational complexity and runtime by eliminating certain inputs.
  - Determine the effect of differing inputs to the load models.
- Determine the feasibility of any real world practical applications for modelling techniques described, with reference to:
  - Time-series forecasting of the expected average load, where short term forecasting accuracy is expected.
  - Predictive calculation of total energy consumption within specific medium term horizons.

6.1.2 Application of ANFIS to model and predict the hot water energy consumption profiles associated with student residences

To satisfy the objective outcomes with confidence required extensive research of all aspects pertaining to the individual components of the problem. These aspects were designated, researched and presented as a comprehensive literature review. The components of the problem are as follows:

- *Sanitary hot water heating:* The energy requirements of sanitary hot water heating installations in different load sectors were researched to obtain an understanding of how

expected behaviour of the users within the context of those sectors influences the behaviour of the electrical load. It was seen that the load associated with residential applications depends on the diurnal and weekly cycles of user behaviour [5], [7]. Also taken into consideration for the scope of the research was of the impact seasonal changes on the target load. Here it was seen that for residential institutes a change in seasons translates to a change in weather patterns, time of the sunrise and environmental factors such as ambient temperature, which all has an effect on user behaviour and consequently the load behaviour [7].

- *Machine learning and ANFIS:* Machine learning can be somewhat of an esoteric field, as the proper implementation of machine learning techniques to specific problems often requires experience and in depth understanding of the workings of those techniques. An understanding of the first principles of the field was garnered through the studying of several textbook sources [27], [25], [21], [26]. This precursor knowledge was vital for the study of the workings of the ANFIS algorithm [10].
- *Matlab fuzzy logic toolbox:* The only real research aid to understanding the design options available for ANFIS training, testing and validation within the fuzzy logic toolbox was in the form of the help documentation compiled by Mathworks on the topic [33].
- *Forecast accuracy metrics:* The error metrics associated with error magnitude calculations as well as the evaluation of time-series predictions were investigated. It was determined that the accuracies of the load forecast models should be presented as Mean Absolute Percentage Error (MAPE) values, due to the prevalence of the use of this metric in research as well as the ease of interpretation of an error that is shown as a percentage [37].
- *Load modelling with ANFIS:* several examples of research for the purpose of STLF were studied in order to hypothesise the necessary inputs required for medium term forecasting. In general the STLF models incorporate inputs such as the number of the half-hourly timeslot of the day, temperature, average load value of the previous three weeks at the same time of the day as prediction, the load one day prior to the same time of day as the prediction and the load 30 minutes prior to the time of prediction [12]. While the timeslot of the day and temperature are viable inputs for consideration for use in medium term predictive models, inputs related to the value of prior loads is not.

The following inputs were hypothesised to be necessary for medium term load approximation of the target data:

- *Time of day:* The possible values of the time of day input are 1 to 48, each indicative of a half-hourly time period.
- *Day of week:* The input is translated to a quantitative numerical equivalent representing the seven days of the week with the numbers 1 to 7, i.e. Monday to Sunday.
- *Day of year:* The numerical qualitative measure of this input ranges from 1 to 366 to allow for leap years.
- *Temperature:* Like the load, the temperature is a quantitative value with a range of 0 to the maximum value in the set.

The load data was pre-processed to remove all days that would perform in a way that is non-characteristic of the normal target load behaviour over similar inputs. Examples of this behaviour include academic reses periods, public holidays that fall on a weekday, second examination periods and outlier data at the beginning and end of semesters during periods of varying occupancy. A single ANFIS utilising all of the descriptive inputs was trained and tested 4 times, each time selecting a different year as testing data and the rest for training, in accordance with the k-fold cross validation method. From the validation procedure it was found that the load model forecasts the expected time-series profile of the load with an average MAPE of 24.6%. Despite what looks like a relatively high time-series prediction error when compared with that obtained by STLF models, the total energy prediction calculated from the time-series profiles exhibits extremely low percentage errors. The 2008 and 2009 calendar years are under predicted by 2.91%, 2010 is over predicted by 2.71% and 2011 by 2.92%. The disparity between the magnitude of the time-series MAPE and the PE of the total energy calculated from the very same profiles suggests that the time-series profile must under- and over predict to a relatively equal amount. Indeed it was shown that all models show a higher tendency for over prediction of low load values and under prediction of high values (See Appendix B).

Analysis of the MAPE over specific timeslots revealed that the model exhibits a much higher rate of error for predictions in the intervals of timeslot 3 to 16, which coincides roughly with the periods of reduced power use associated with the early mornings. There is a second period of increased error rate in the evening between timeslot 30 and 37, although far less pronounced, which coincides with the traditional periods of reduced power use seen between the morning and evening peak. It was concluded that there are two reasons for this consistent spike in the MAPE over the daily timeslots:

- The daily output profile of the predicted load is “smoothed” out when compared with the actual data of the same period. While the observed load exhibits virtually no load in the early morning followed by a rapid ramp up to morning peak values at roughly 7:30, the predicted profile exhibits a gradual change from roughly 6:00 to 7:30. The “smoothness” of the predicted profile due to the use of a single ANFIS model hinders the prediction of the rapid changes in power use that occur for real load profiles.
- A second contribution to the poor MAPE performance of the model was found to be an unfortunate shortcoming of the MAPE accuracy metric when applied to predictions of very low values. A relatively small prediction error of a near zero value can easily be many times the actual size of the value, resulting in a very large percentage error. This large error during these times may not be a problem when evaluating a load profile with a continuous baseload.

Overall, it is concluded that it is feasible to model and predict the hot water energy consumption profiles associated with student residences even with a simple single model implementation of the ANFIS method utilising a high number of inputs. This finding serves as a good starting platform for the next research objective, i.e. searching for a more optimal model configuration through different permutations of inputs and sub-model compartmentalisation.

6.1.3 Comparative case study of different ANFIS implementations

6.1.3.1 Overview

There is flexibility with regards to the combinations of inputs which can be used, since both the day of year input and the temperature input should be able to supply the necessary information to infer the time of the year to an ANFIS model. Furthermore a model can be compartmentalised into sub-models that either target each weekday category or each timeslot of the day separately, thereby reducing the number of training inputs to the model.

A case study structure is devised that consists of four groups, namely case groups A, B, C and D. Each group differs in that a different setup of inputs is made available to each. Group A is trained with all inputs, group B with all but the temperature, group C is trained with all but the day of year input and group D is trained with all the inputs, but with a 30 minute delay on the temperature input. The first case in each group consists of a single ANFIS model targeting all of the inputs made available to that group. The second consists of three sub-models that each

targets a weekday category and only uses the weekday input to correctly index data to those sub-models. The third does the same with 48 sub-models targeting distinct times of the day, and the fourth (though only possible for groups A and D) targets both time of day and weekday categories. Through comparisons between the case study results, conclusions could be drawn of the effects of different inputs and different sub-model architectures.

All in all the fuzzy-logic toolbox only supplies the user with functions to generate, train and test individual models with individual sets of training and testing data. Extensive functionality was written into a Matlab class structures to accommodate the partitioning, indexing, training, testing and validation of many different sub models in various configurations. Additionally, comprehensive graphing functionality was written to enable the display of the observed- and predicted load and temperature profiles in the form of both 3D mesh plots and heat map renderings.

6.1.3.2 Load model partitioning

It was found that partitioning of the load models into the proposed sub-models had the following effects for each method of model compartmentalization:

- *Sub-models for weekdays:* The effect of this method was consistent across all case groups, and saw a very slight increase in prediction accuracy from the single-model systems. This was primarily the result of slightly increased accuracies for weekday predictions, with no improvements seen for weekend predictions.
- *Sub-models for time of day periods:* This proved to be the most successful compartmentalisation method in each group. Load models A3 and D3 have practically the same MAPE of 21%, model B3 is only slightly worse at 21.61%, and Model C3 at 23.91%. The lower MAPE is no doubt due to the increased ability of the model to adjust to a rapid ramp rate of observed data for the morning peak.
- *Sub-models for both of the above:* This method gave good accuracy results for weekdays, even comparable to the performances of A3 and D3 in that regard. However they were found to perform very poorly for weekends, as the training dataset is compartmentalised to such a degree that the data available for weekend training is very limited.

Furthermore it was found that the load model runtime is a function of the number of model inputs. The mean training runtime for case models is roughly 50 minutes with 4 inputs, 10 minutes with 3 inputs, and 2 to 5 minutes with 2 inputs. Therefore a reduction in the number of inputs has a notable effect on the computational complexity.

6.1.3.3 Differing inputs

The effects of the different inputs made available to each case group is as follows:

- *Group A:* benchmark group with all inputs.
- *Group B:* Omission of the temperature variable seems to not have a negative effect on either the time-series or total energy prediction accuracies when compared to the counterpart models in group A. With the minor exception being model A3, which performs slightly better that B3.
- *Group C:* While unquestionably the worst performing group in the set, it is interesting to note that model C3 still outperforms model A1. This attests to the usefulness of model compartmentalisation and the ability of the ANFIS models to infer the general time of year from nothing but the temperature input.
- *Group D:* The delayed temperature input has no noticeable effect on the performance of the group models when compared to group A.

6.1.4 Feasibility for practical applications

6.1.4.1 Time series forecasting of load profiles

It was found that the time-series accuracies of the load models can be improved from the base model through the implementation of a sub-model architecture targeting time periods of the day. It was also pointed out that the MAPE metric is adversely influenced because the target load approaches zero during off peak times. Therefore it is concluded that this method of medium term time-series prediction could be made feasible if used with load data that has a persistent baseload.

6.1.4.2 Calculation of cumulative energy consumption

It was shown that almost every model that was considered performed equally well when predicting the total energy consumption over a specific range. Each consistently under predicted for years 2008 and 2009, and over predicted for 2010 and 2011. It is concluded that this consistent under-, and over-estimation of specific years by all load models could at least be partially attributed to the yearly decrease in the average energy consumption from 2008 to 2010. This would make this method ideal for measurement and verification baseline adjustments.

6.2 Recommendations for future work

This study gives rise to a number of recommendations for further work, including the following:

- Determine the accuracy measure of the predictions of each load model over seasonal periods, additional to the timeslot and day of week metrics already shown, so that the seasons that exhibit the best and worst behaviours can be identified.
- Perform the presented case study evaluations on other types of load data with the following properties:
  - Constant mean energy consumption over the range of years, so that the calculation of cumulative energy consumption can be better critiqued.
  - Has a consistent baseload, so that the medium term time series forecasting accuracy can be better critiqued.
- Test the methodology on any number of load categories that show seasonal, weekly and daily cyclic patterns.

Bibliography

[1] J. J. Cárdenas, L. Romeral, A. Garcia, and F. Andrade, "Load forecasting framework of electricity consumptions for an Intelligent Energy Management System in the user-side," *Expert Syst. Appl.*, vol. 39, no. 5, pp. 5557–5565, 2012.

[2] C. Barteczko-Hibbert, M. Gillott, and G. Kendall, "An artificial neural network for predicting domestic hot water characteristics," *Int. J. Low-Carbon Technol.*, vol. 4, no. 2, pp. 112–119, 2009.

[3] C. Sigauke and Delson Chikobvu, "Daily peak electricity load forecasting in South Africa using a multivariate non-parametric regression approach," *Orion*, vol. 26, no. 2, pp. 97–111, 2010.

[4] *M & V Standard for 12-L Tax Incentives, SABS : SANS 50010*. 2011.

[5] J. P. Meyer and M. Tshimankinda, "Domestic hot water consumption by developing communities in South African traditional houses," *Energy*, vol. 21, no. 12, pp. 1101–1106, 1996.

[6] R. Rankin and P. G. Rousseau, "Sanitary hot water consumption patterns in commercial and industrial sectors in South Africa: Impact on heating system design," *Energy Convers. Manag.*, vol. 47, no. 6, pp. 687–701, 2006.

[7] J. E. Calmeyer and G. J. Delport, "The modelling and control of hot water consumption in residential hostels," *1999 IEEE Africon. 5th Africon Conf. Africa (Cat. No.99CH36342)*, vol. 2, pp. 825–830, 1999.

[8] R. Joyeux and R. D. Ripple, "Household energy consumption versus income and relative standard of living: A panel approach," *Energy Policy*, vol. 35, no. 1, pp. 50–60, 2007.

[9] E. Valor, V. Meneu, and V. Caselles, "Daily Air Temperature and Electricity Load in Spain," *J. Appl. Meteorol.*, vol. 40, no. 8, pp. 1413–1421, 2001.

[10] J. S. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Trans. Syst. Man Cybern.*, vol. 23, no. 3, pp. 665–685, 1993.

[11] M. Mordjaoui and B. Boudjema, "Forecasting and Modelling Electricity Demand Using Anfis Predictor," *J. Math. Stat.*, vol. 7, no. 4, pp. 275–281, 2011.

[12] W. Yuill, "Application of Adaptive Neuro Fuzzy Inference System (ANFIS) based Short Term Load Forecasting in South African Power Networks," *Expert Syst.*, 2010.

[13] J. D. Glover, M. S. Sarma, and T. Overbye, *Power System Analysis and Design*, vol. 5, no. 7. 2008.

[14] G. M. Masters, *Renewable and Efficient Electric Power Systems*. 2004.

[15] B. Tarroja, F. Mueller, J. D. Eichman, and S. Samuelsen, "Metrics for evaluating the impacts of intermittent renewable generation on utility load-balancing," *Energy*, vol. 42, no. 1, pp. 546–562, 2012.

[16] W. L. . den Heijer, "The Measurement and Verification Guideline for Energy Efficiency and Demand-Side Management (EEDSM) Projects and Programmes," Potchefstroom, 2010.

[17] Eskom, "Switching off your geyser helps everybody," 2013. .

[18] M. Beccali, M. Cellura, V. Lo Brano, and a. Marvuglia, "Forecasting daily urban electric load profiles using artificial neural networks," *Energy Convers. Manag.*, vol. 45, no. 18–19, pp. 2879–2900, 2004.

[19] P. P. Bonissone, "Soft computing: the convergence of emerging reasoning technologies," *Soft Comput. - A Fusion Found. Methodol. Appl.*, vol. 1, no. 1, pp. 6–18, 1997.

[20] a. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, 1999.

[21] V. Cherkassky and F. Mulier, *Learning from Data: Concepts, Theory, and Methods*, vol. 43, no. 1. 1998.

[22] R. Rankin, P. G. Rousseau, and M. Van Eldik, "Demand side management for commercial buildings using an inline heat pump water heating methodology," *Energy Convers. Manag.*, vol. 45, no. 9–10, pp. 1553–1563, 2004.

[23] G. L. Morrison, T. Anderson, and M. Behnia, "Seasonal performance rating of heat pump water heaters," *Sol. Energy*, vol. 76, no. 1–3, pp. 147–152, 2004.

[24] A. Hepbasli and Y. Kalinci, "A review of heat pump water heating systems," *Renew. Sustain. Energy Rev.*, vol. 13, no. 6–7, pp. 1211–1229, 2009.

[25] T. M. Mitchell, *Machine Learning*, vol. 4, no. 1. 1997.

[26] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Second Edi., vol. 1. Springer, 2009.

[27] J. Lawrence, *Introduction to Neural Networks: Design, Theory, and Applications*, Fifth edit. Nevada: California Scientific Software Press, 1993.

[28] T. Dietterich, "Limitations on inductive learning (extended abstract)," in *proceedings of the sixth international workshop on machine learning*, 1989, pp. 124–128.

[29] L. a. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338–353, 1965.

[30] J.-S. R. Jang, "Input selection for ANFIS learning," *Proc. IEEE 5th Int. Fuzzy Syst.*, vol. 2, pp. 1493–1499, 1996.

[31] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *Syst. Man Cybern. IEEE Trans.*, vol. SMC-15, no. 1, pp. 116–132, 1985.

[32] J. S. R. Jang and C. T. Sun, "Neuro-Fuzzy Modeling and Control," *Proc. IEEE*, vol. 83, no. 3, pp. 378–406, 1995.

[33] MathWorks, "Fuzzy Logic Toolbox Documentation," 2015. .

[34] C. Tofallis, "A better measure of relative prediction accuracy for model selection and model estimation," *J. Oper. Res. Soc.*, vol. 66, no. 8, pp. 1352–1362, 2015.

[35] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature," *Geosci. Model Dev.*, vol. 7, no. 3, pp. 1247–1250, 2014.

[36] D. Montgomery and R. Runger, *Applied statistics and probability for engineers*. Hoboken: John Wiley and Sons, Inc, 2011.

[37] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE)

over the root mean square error (RMSE) in assessing average model performance," *Clim. Res.*, vol. 30, no. 1, pp. 79–82, 2005.

[38] L. Ying, "Using adaptive network based fuzzy inference system to forecast regional electricity loads," vol. 49, pp. 205–211, 2008.

[39] P. Kumar, "Minimum Weekly Temperature Forecasting using ANFIS w + w," vol. 3, no. 5, pp. 1–6, 2012.

[40] S. Kaynak, H. Evirgen, and B. Kaynak, "Adaptive Neuro-Fuzzy Inference System in Predicting the Success of Student's in a Particular Course," *Int. J. Comput. Theory Eng.*, vol. 7, no. 1, pp. 34–39, 2014.

Appendix A Model Outputs for Cross Validation Training

Appendix A.1 Model A1

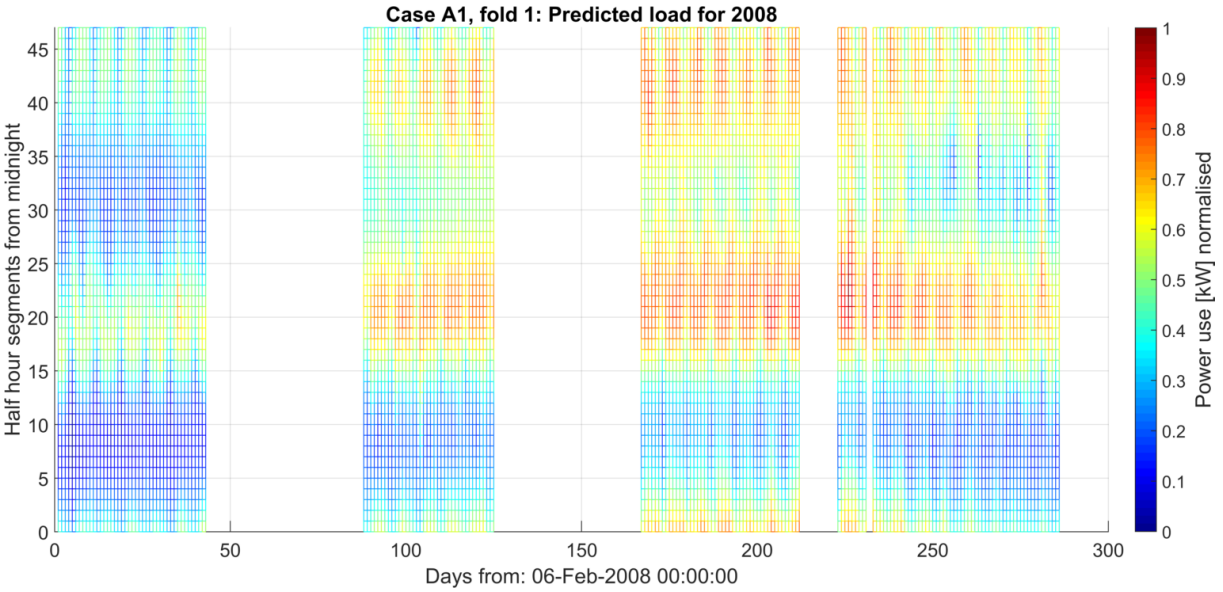

Figure A-1: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model A1.

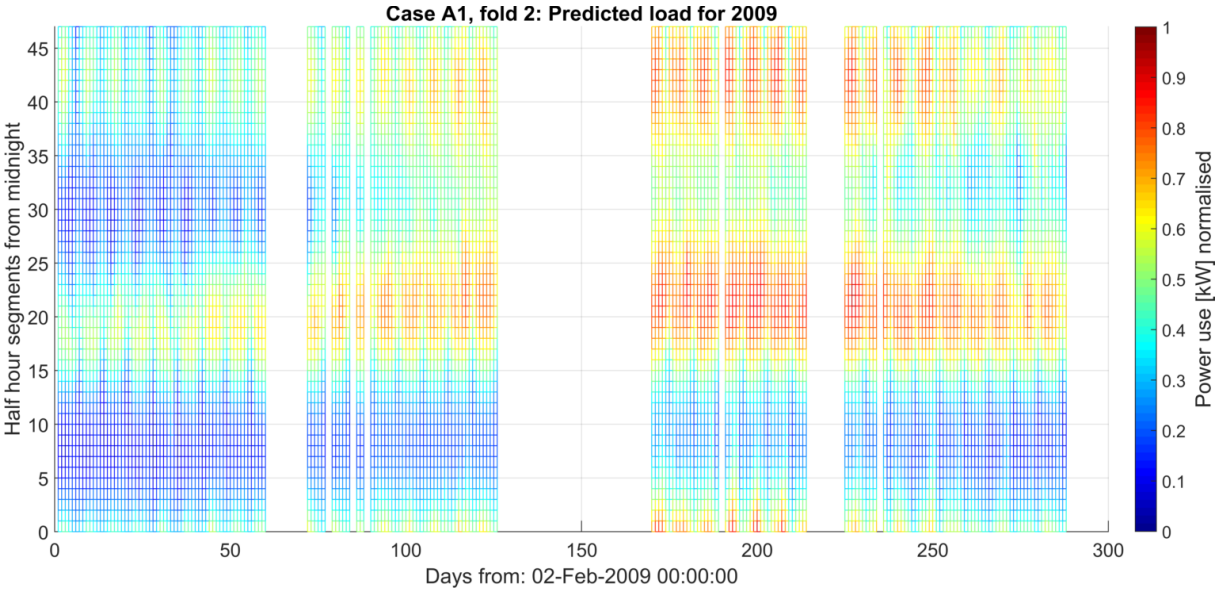

Figure A-2: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model A1.

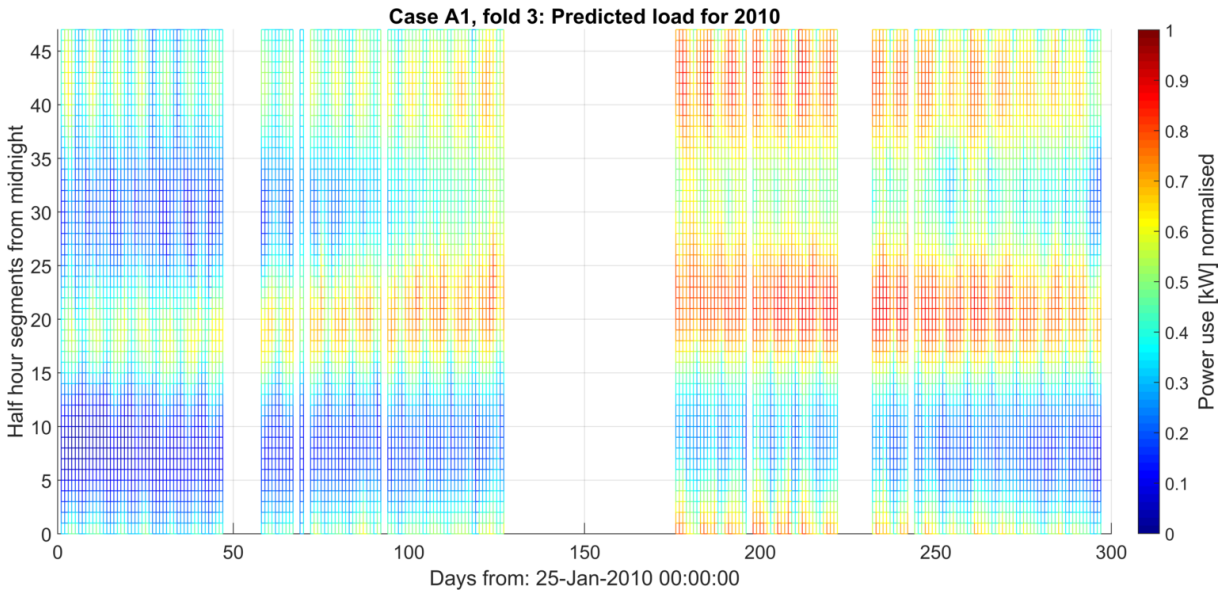

Figure A-3: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model A1.

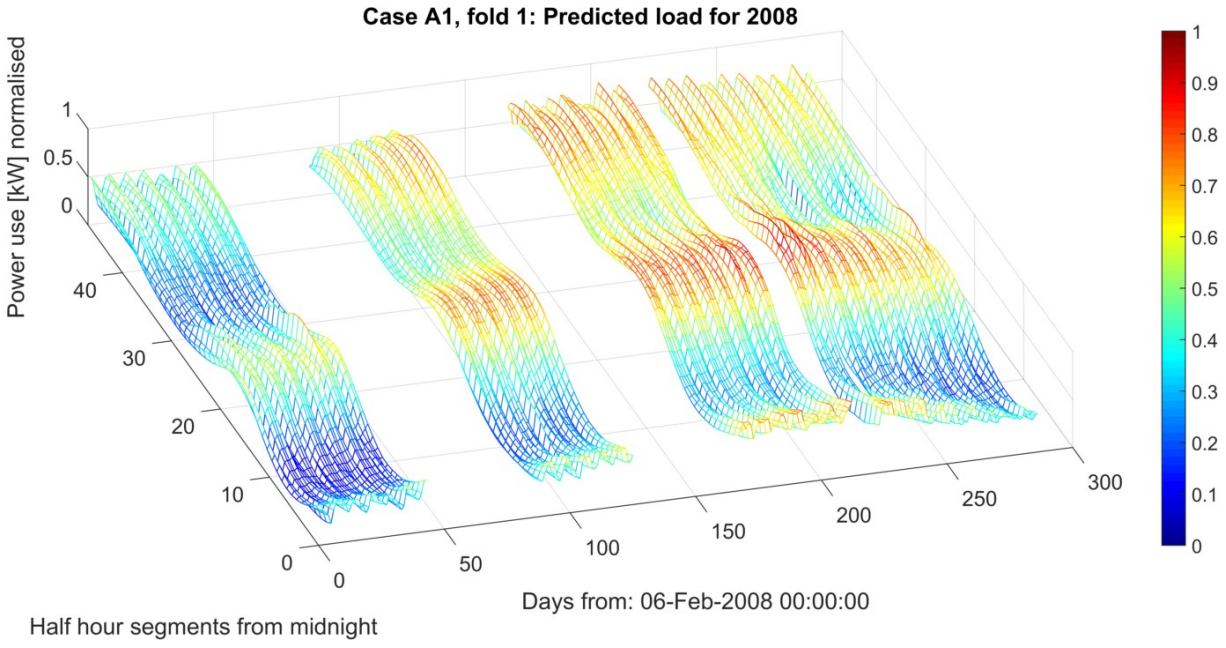

Figure A-4: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model A1.

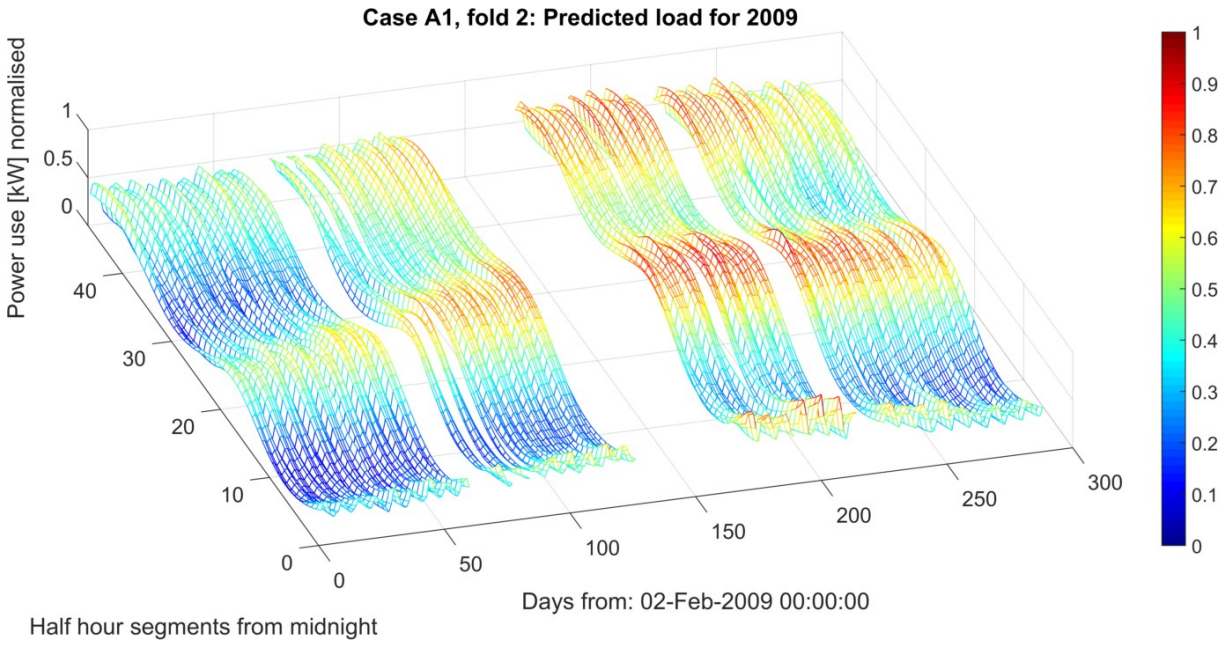

Figure A-5: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model A1.

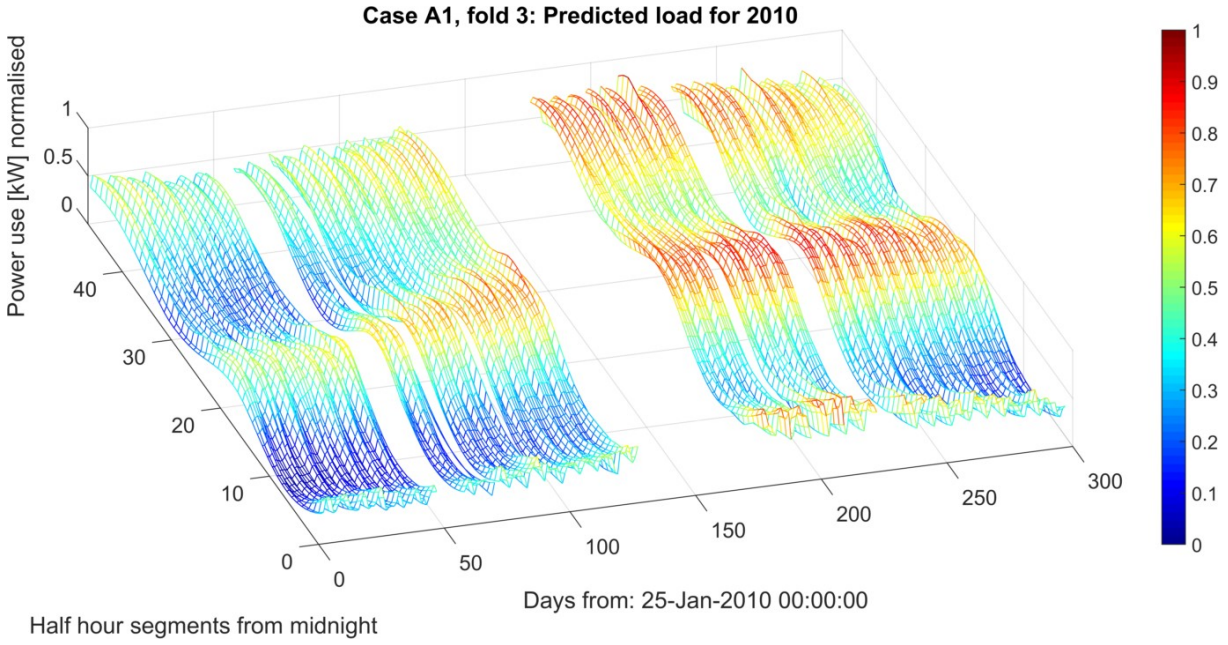

Figure A-6: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model A1.

Appendix A.2 Model A2

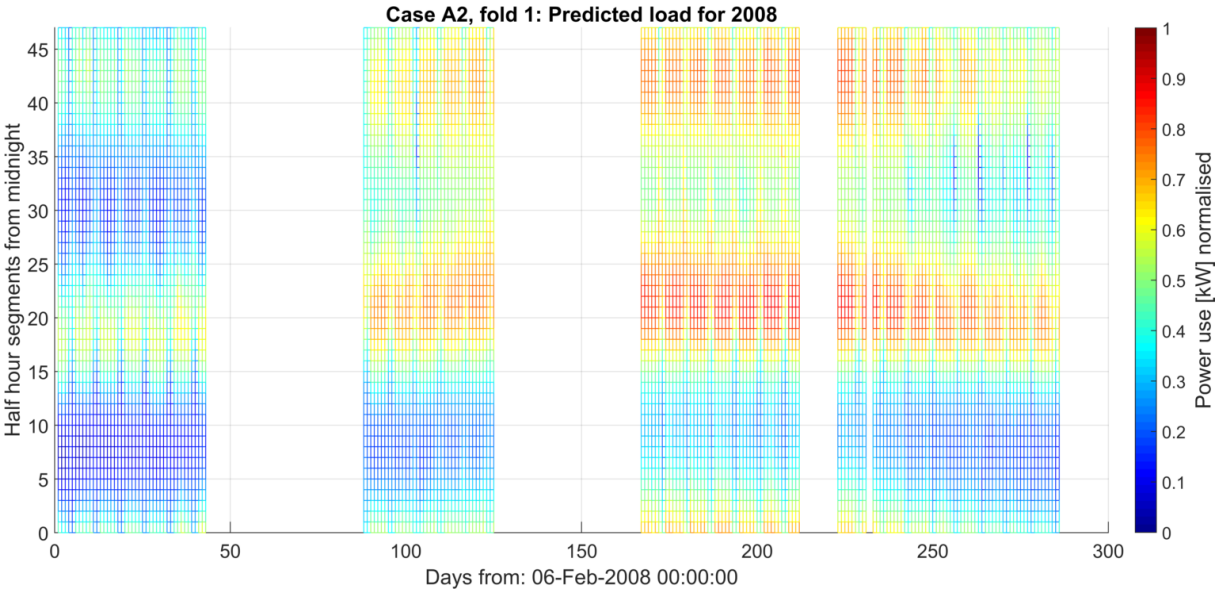

Figure A-7: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model A2.

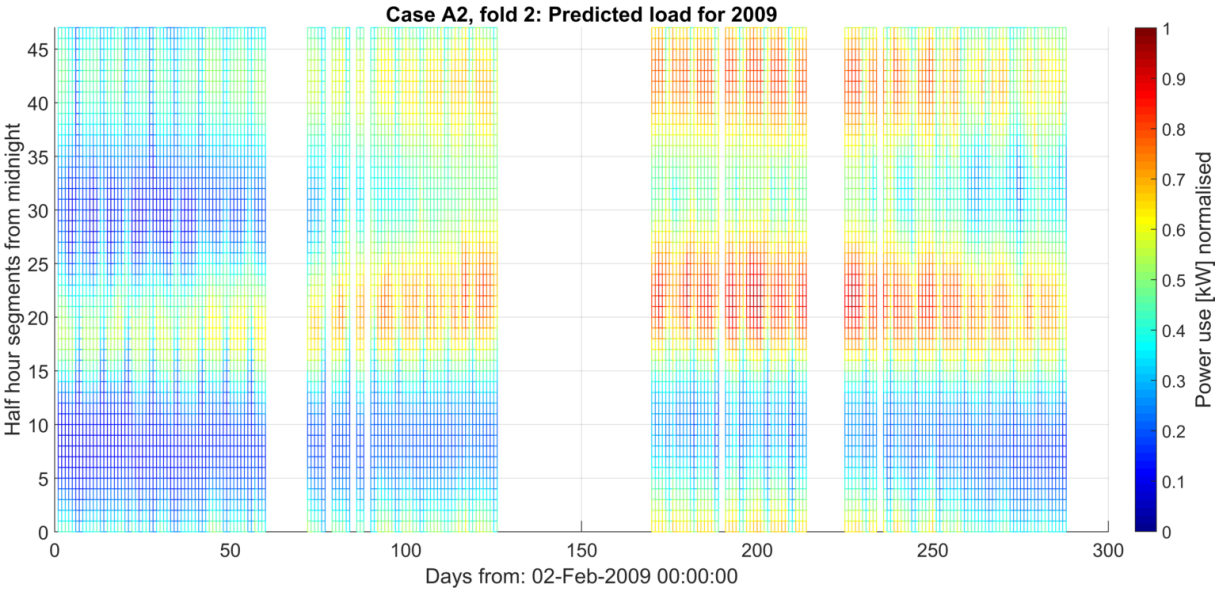

Figure A-8: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model A2.

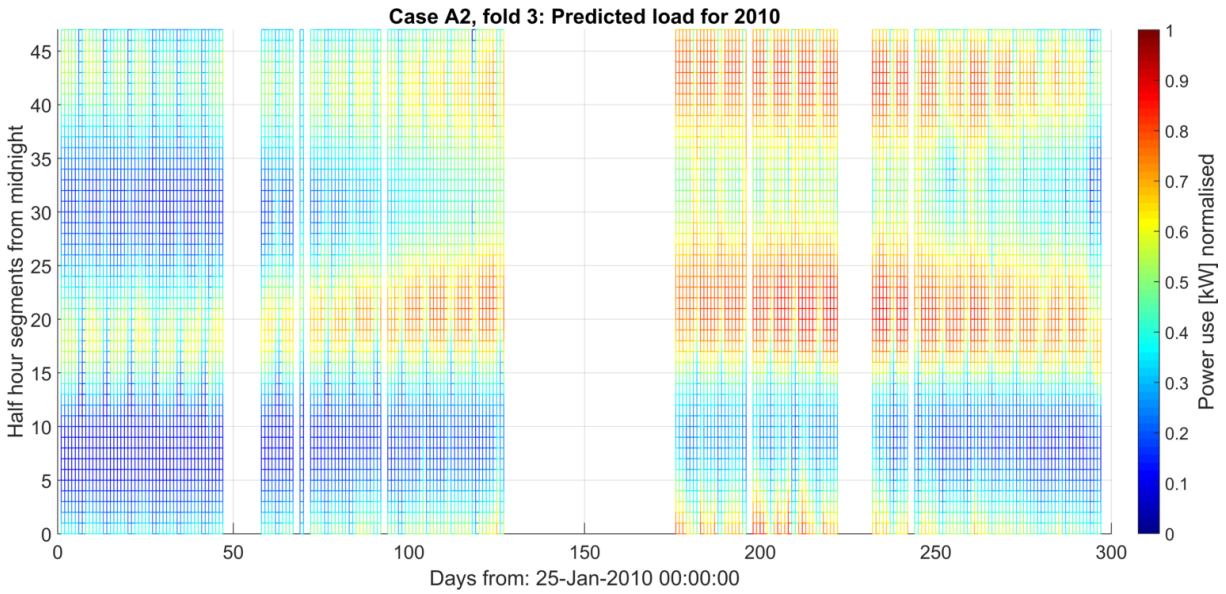

Figure A-9: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model A2.

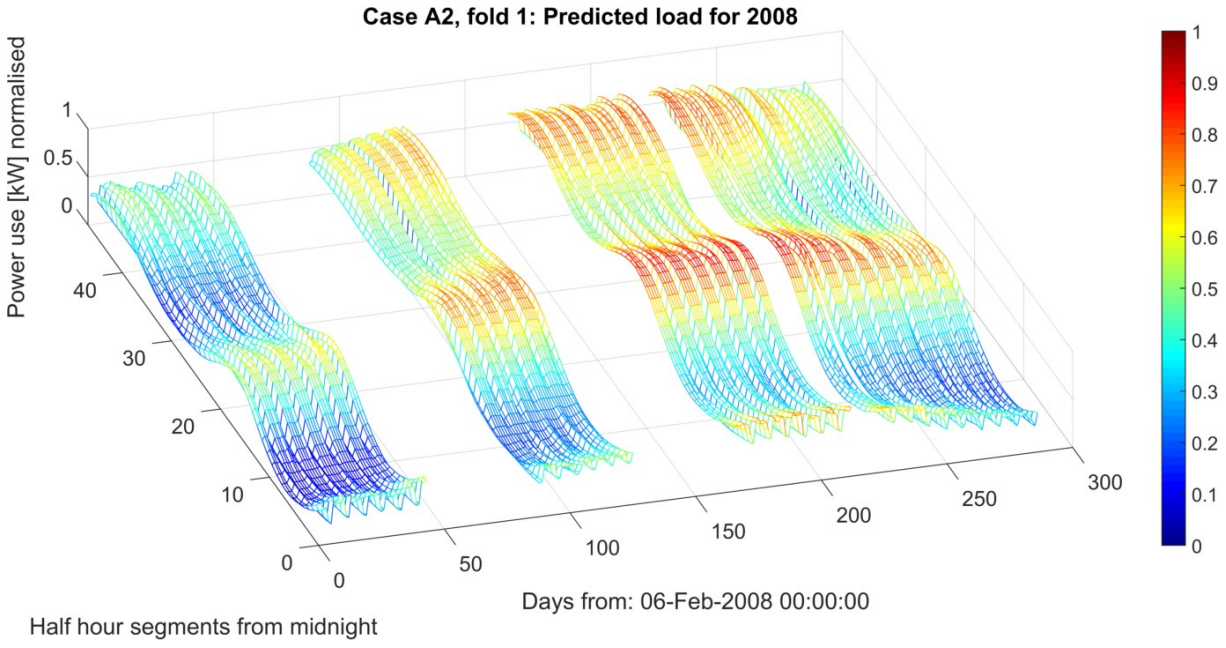

Figure A-10: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model A2.

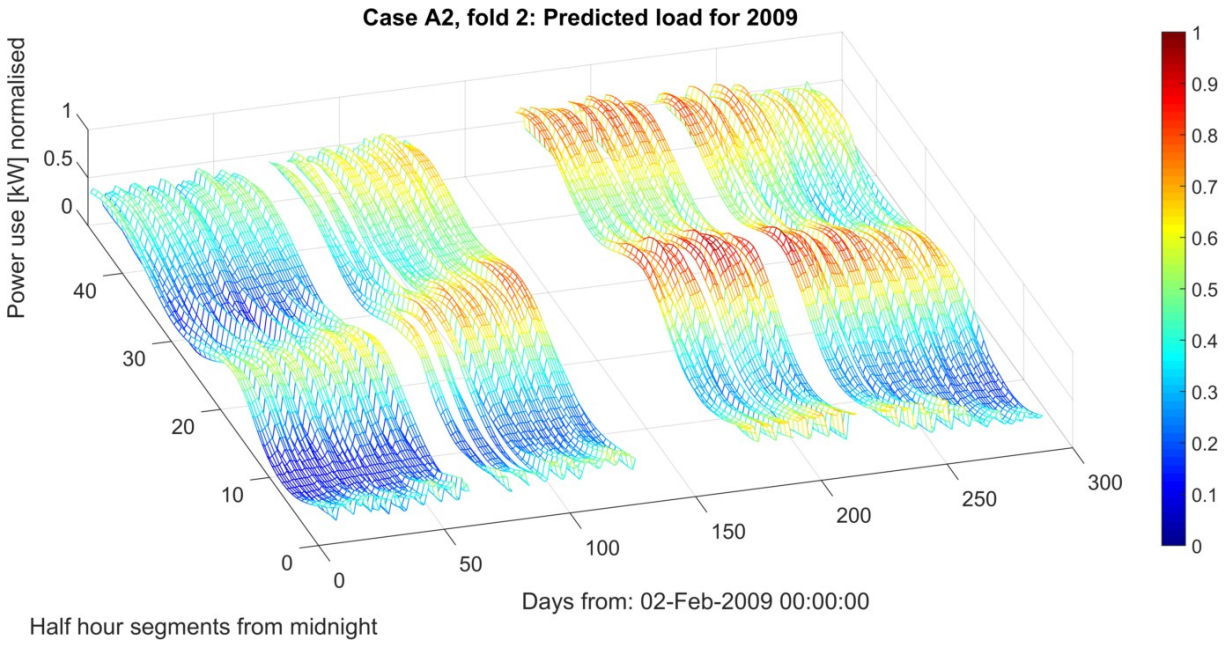

Figure A-11: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model A2.

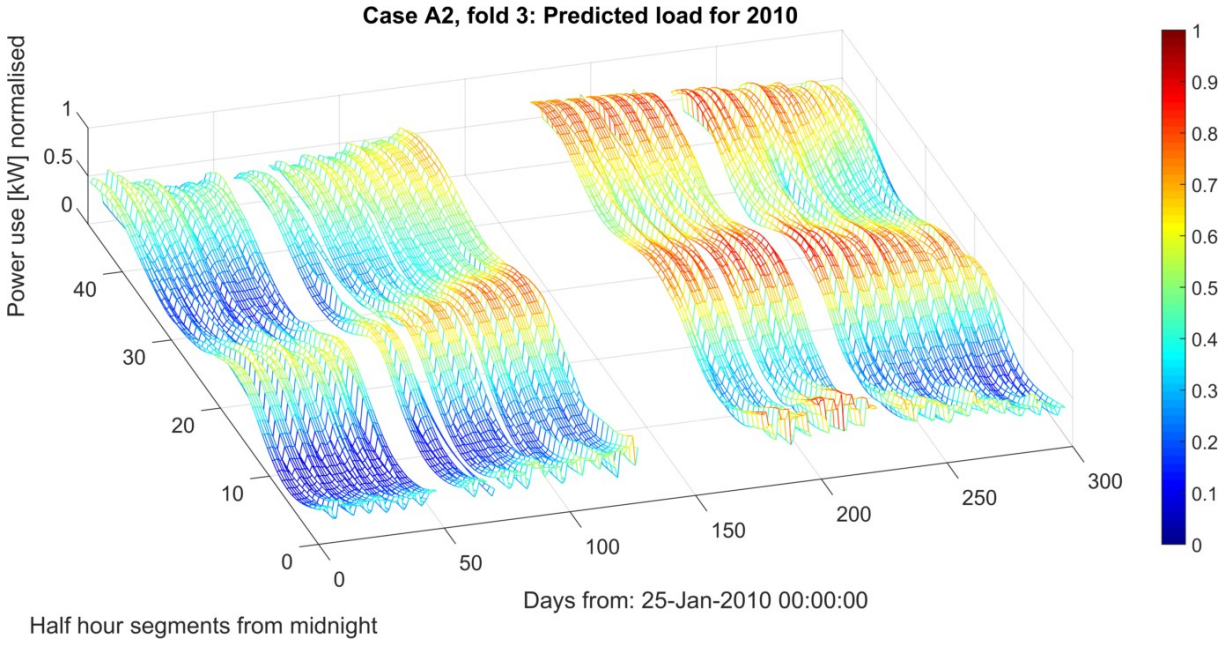

Figure A-12: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model A2.

Appendix A.3 Model A3

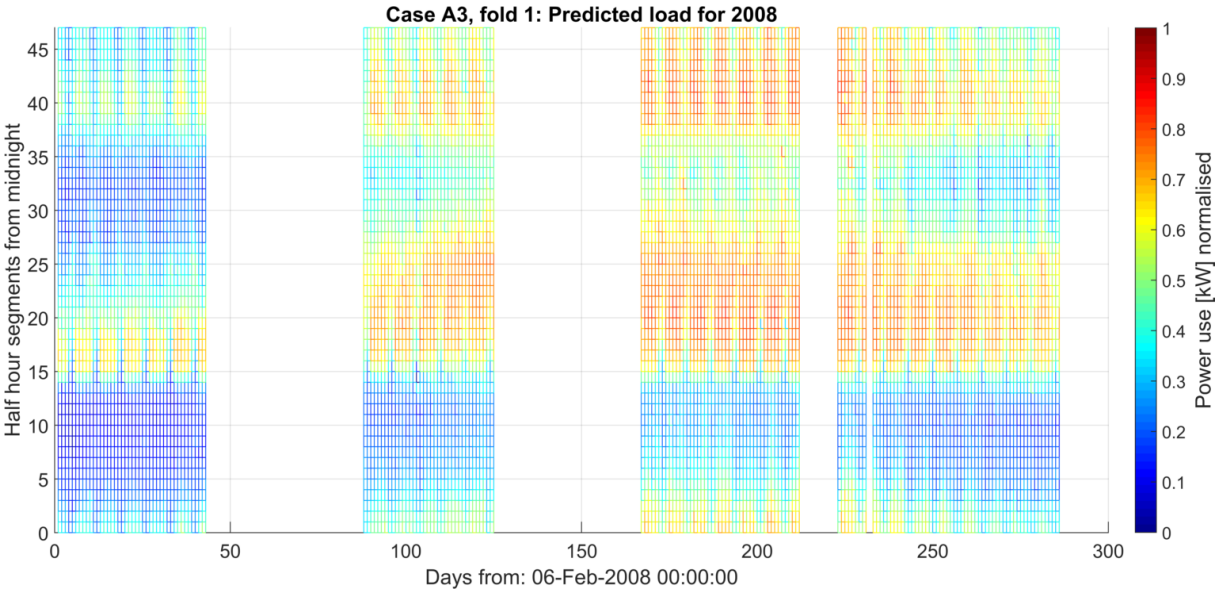

Figure A-13: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model A3.

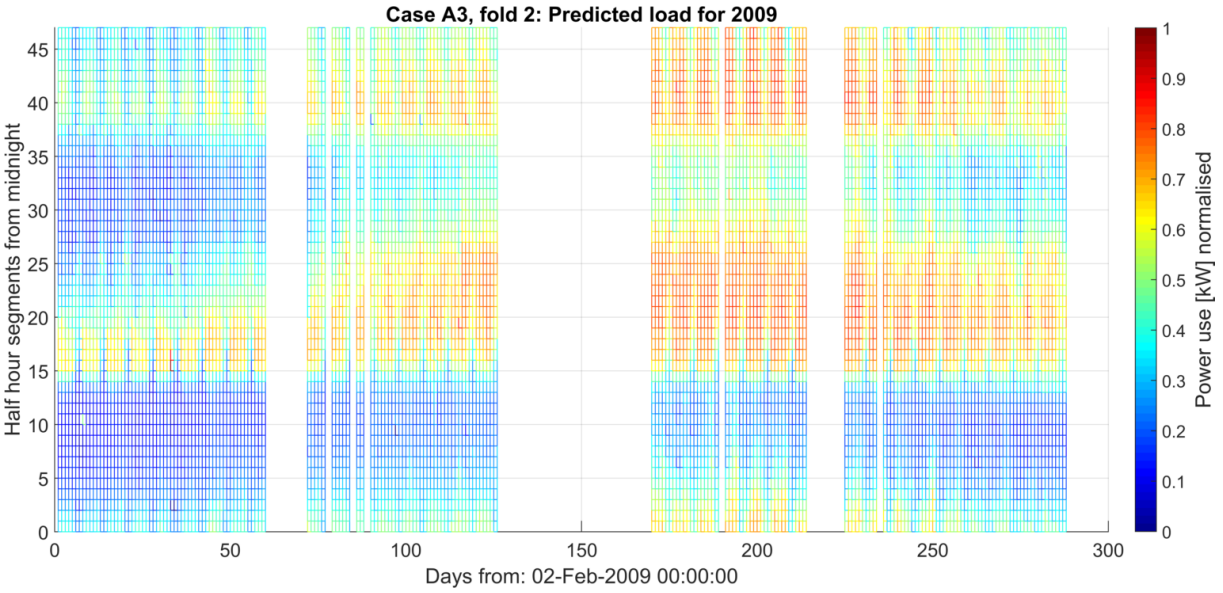

Figure A-14: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model A3.

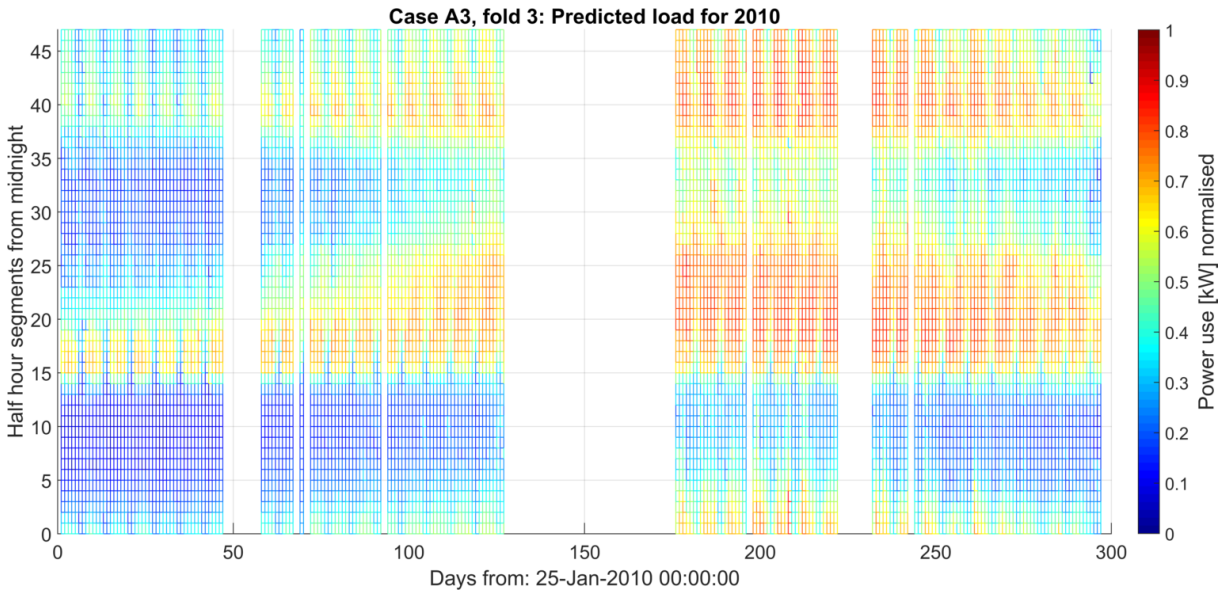

Figure A-15: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model A3.

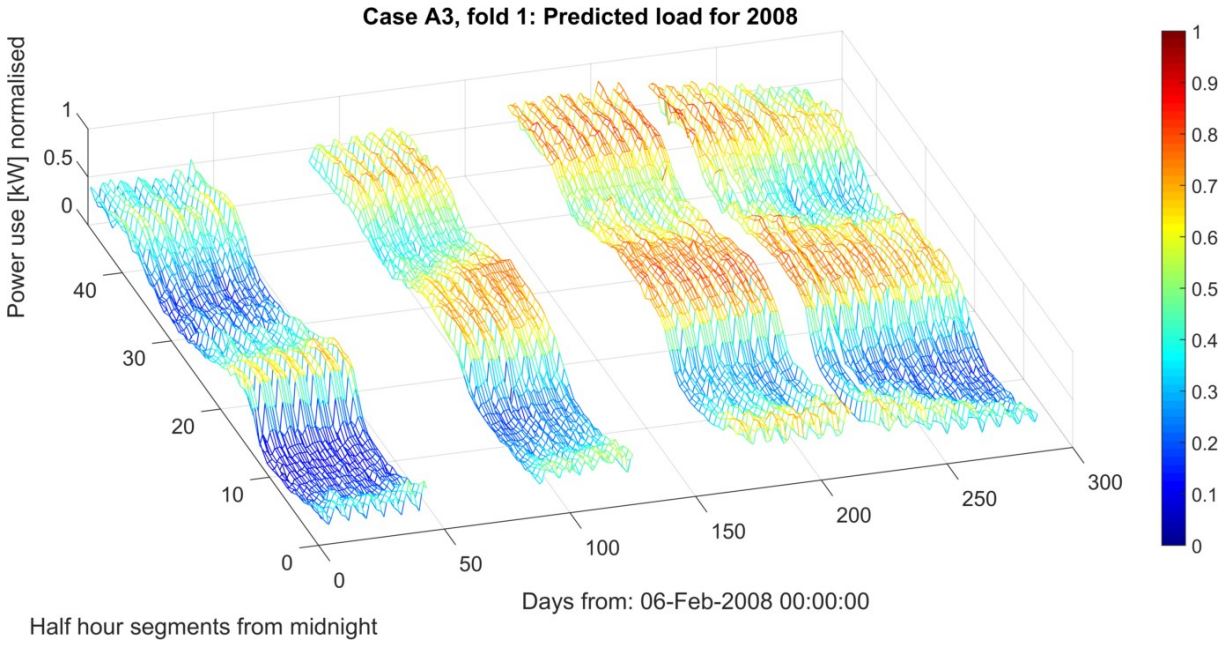

Figure A-16: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model A3.

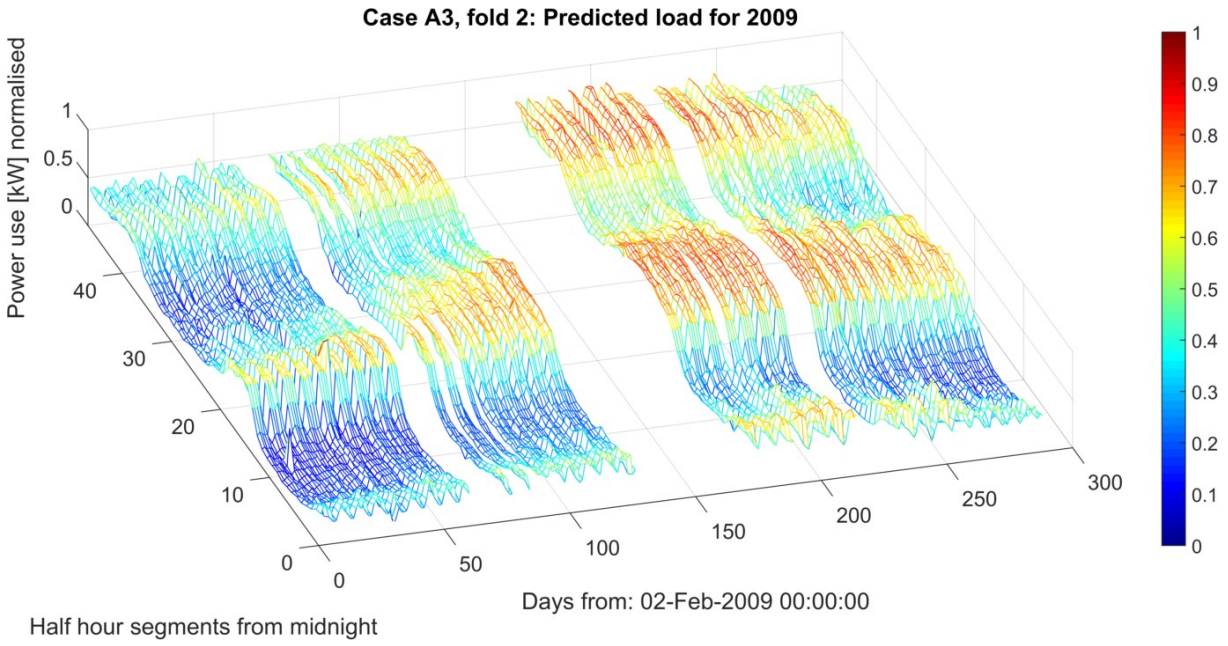

Figure A-17: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model A3.

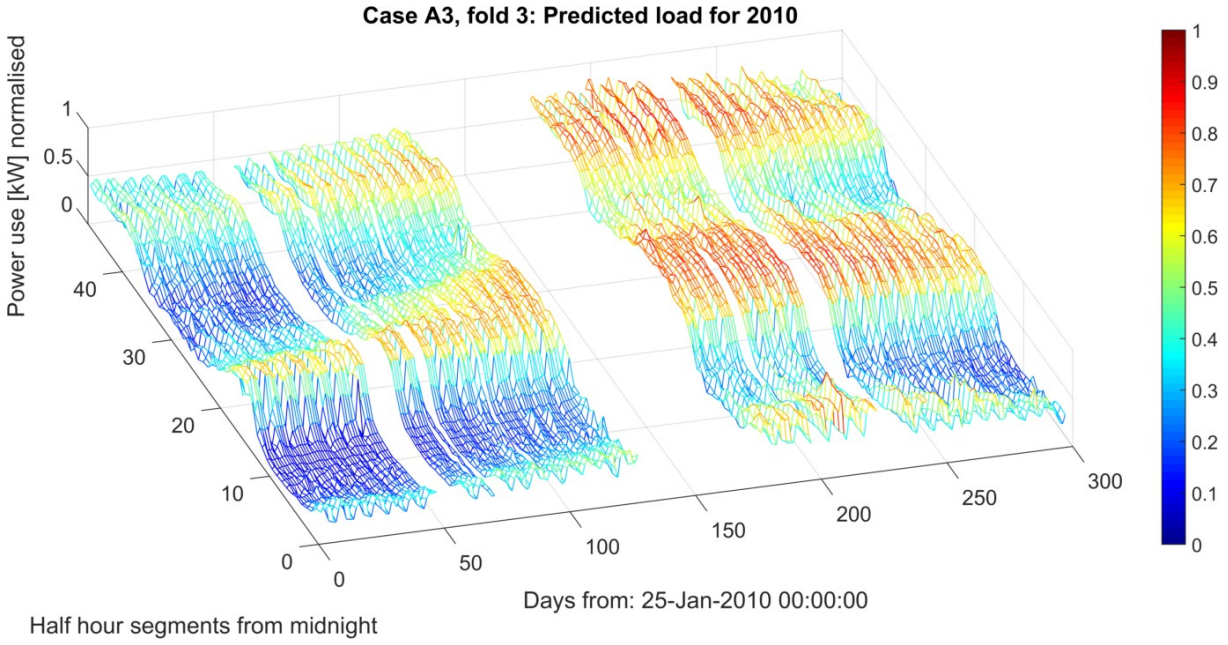

Figure A-18: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model A3.

Appendix A.4 Model A4

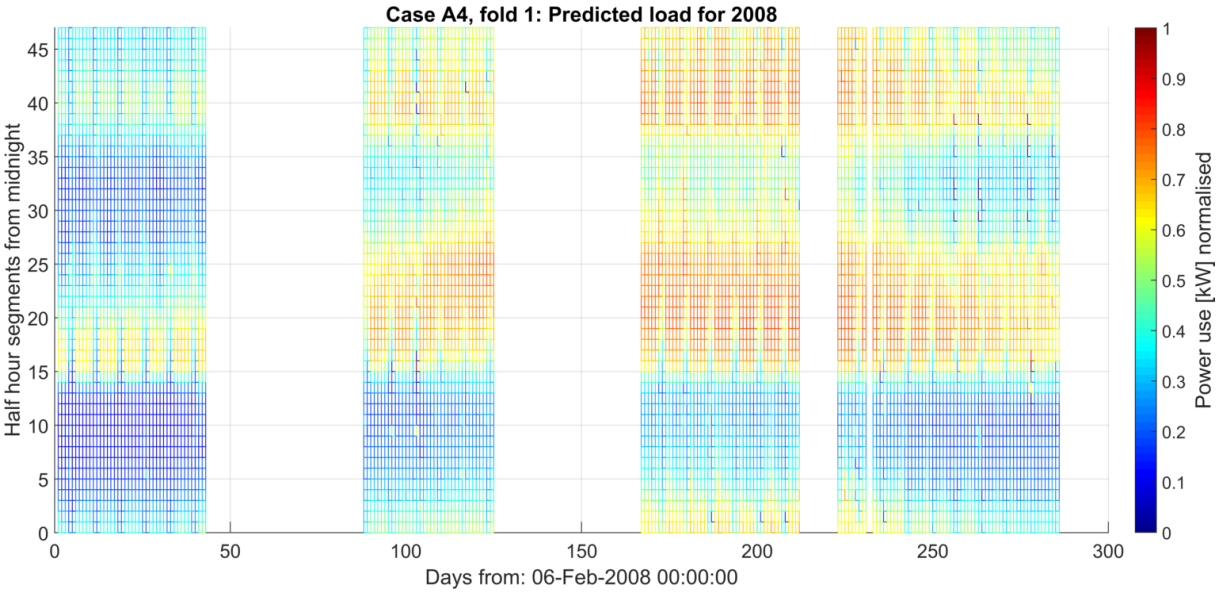

Figure A-19: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model A4.

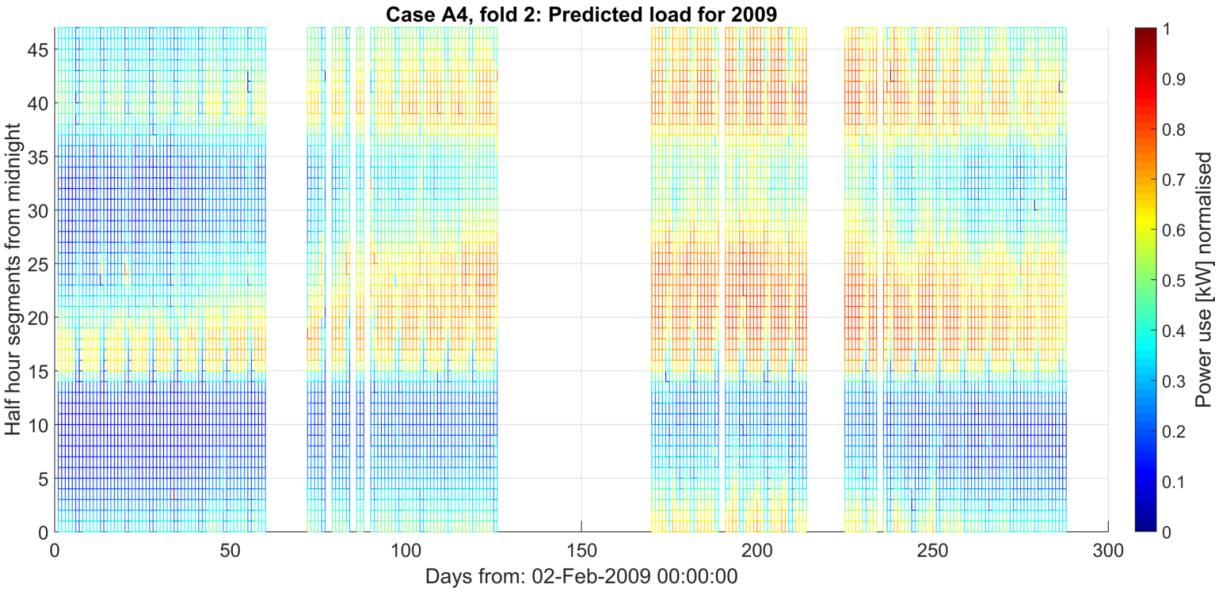

Figure A-20: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model A4.

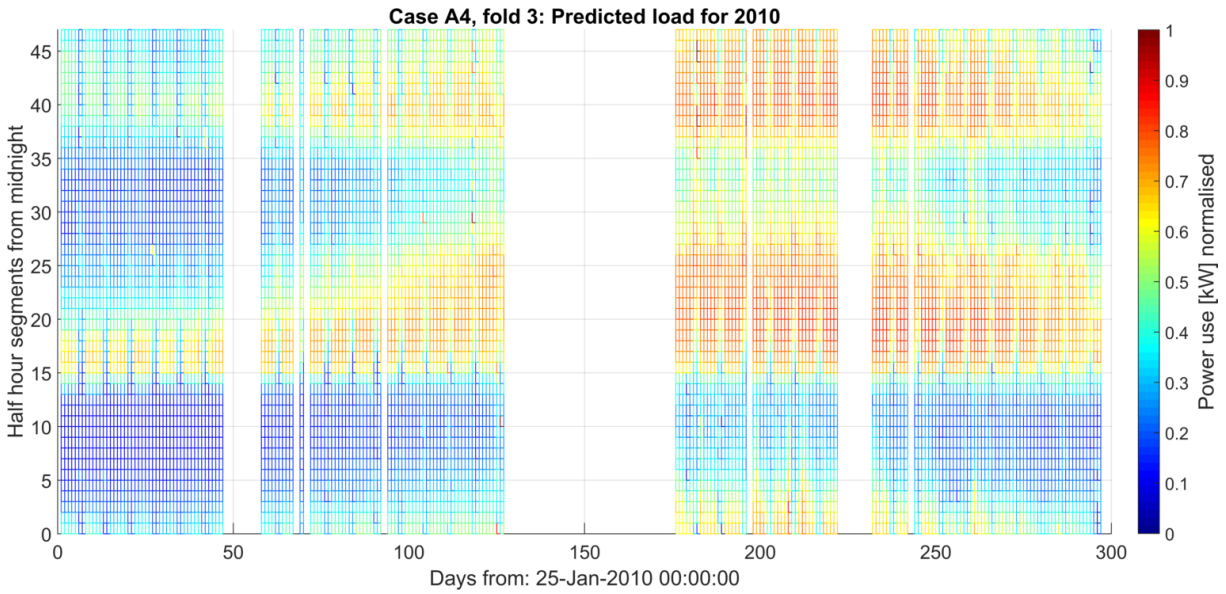

Figure A-21: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model A4.

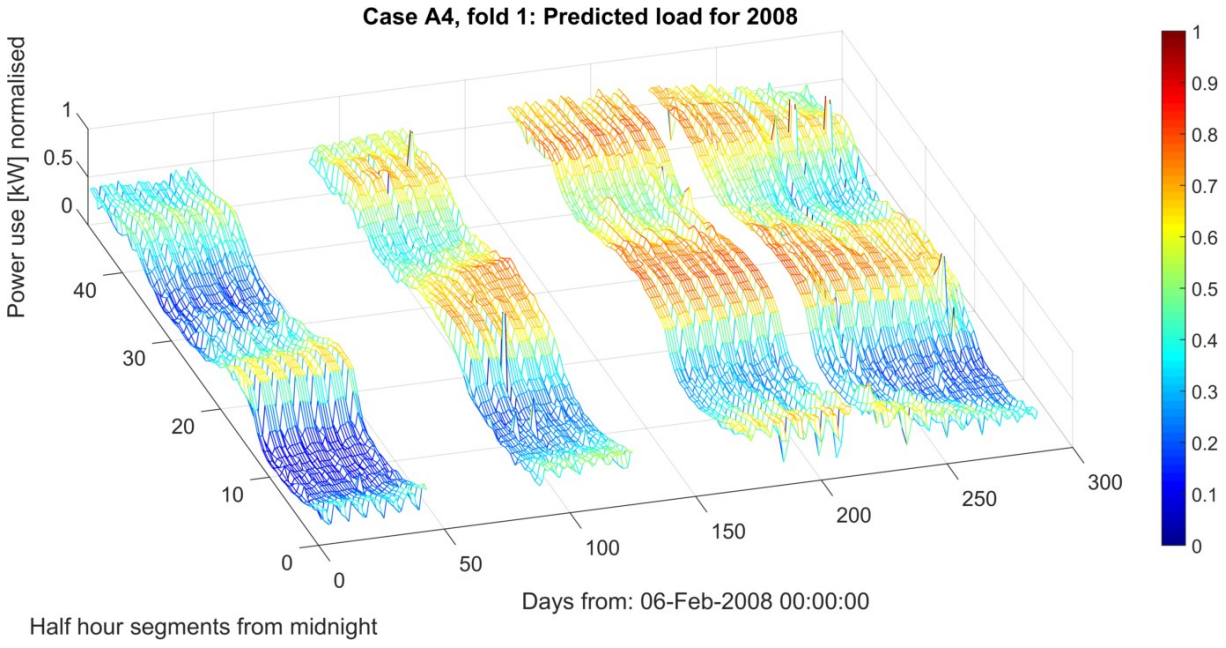

Figure A-22: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model A4.

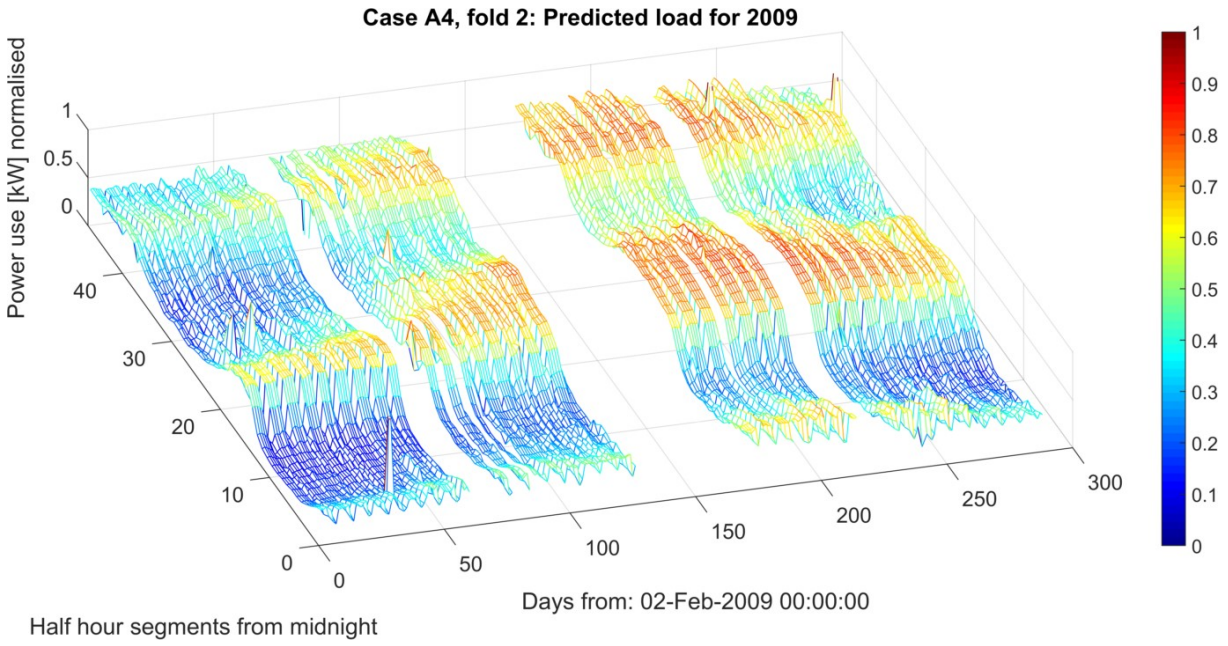

Figure A-23: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model A4.

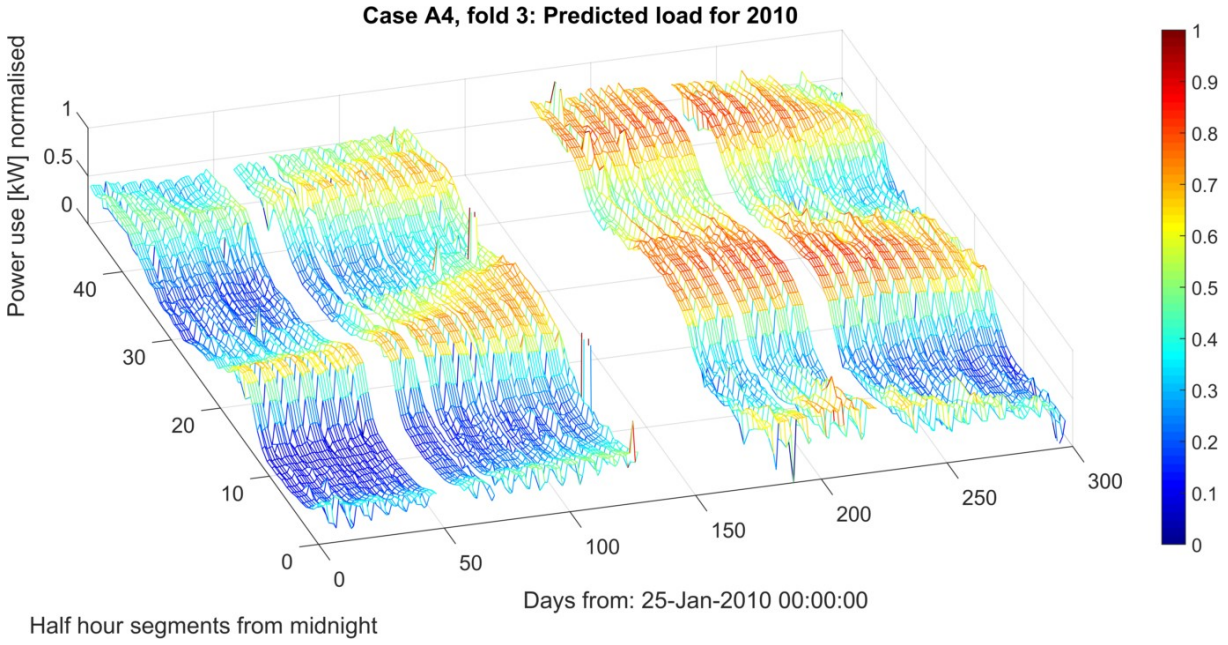

Figure A-24: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model A4.

Appendix A.5 Model B1

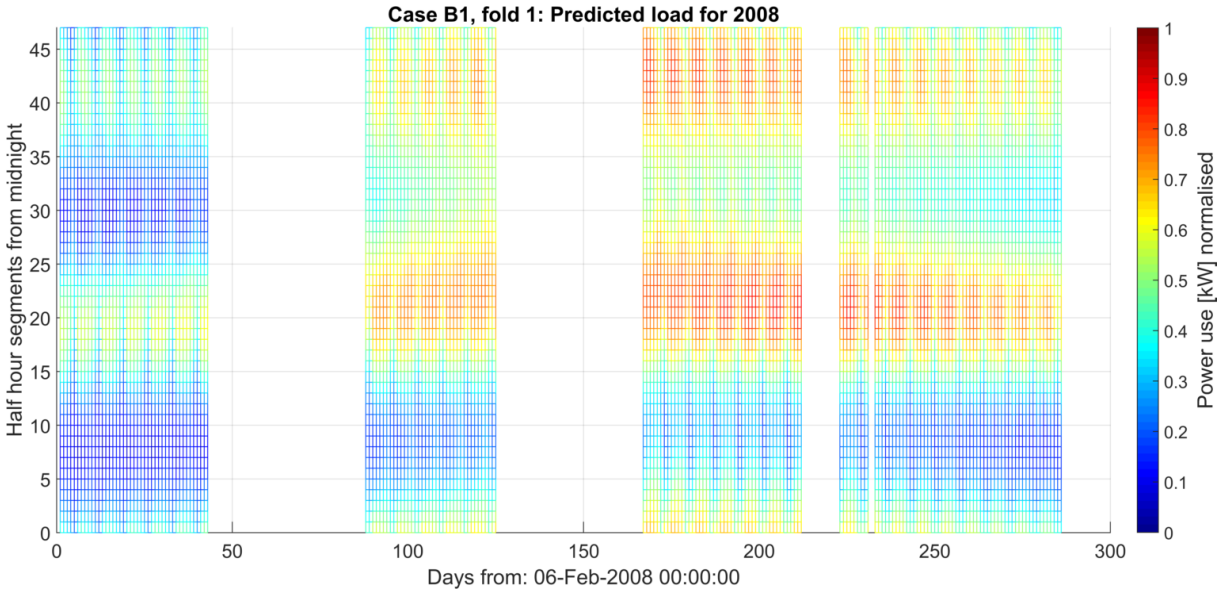

Figure A-25: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model B1.

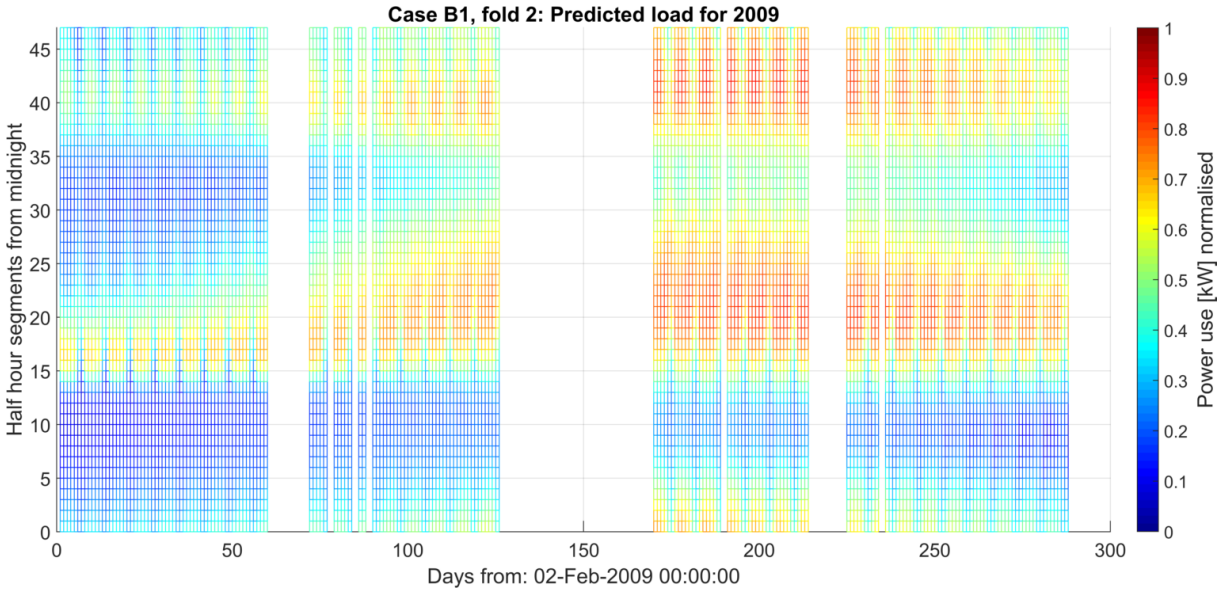

Figure A-26: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model B1.

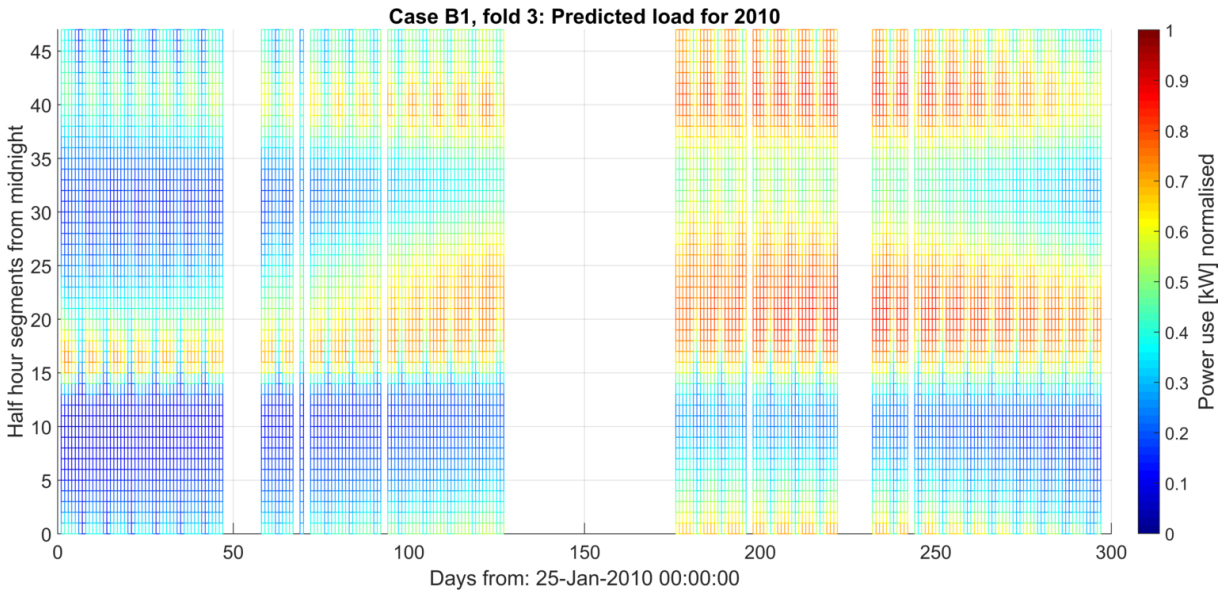

Figure A-27: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model B1.

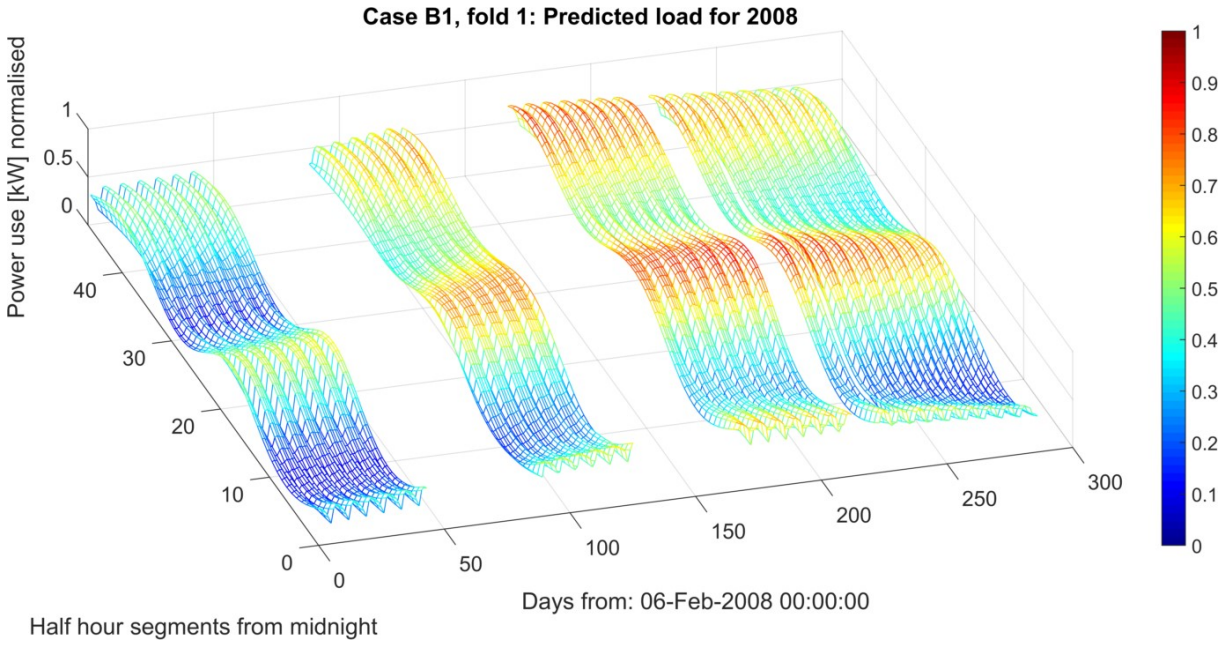

Figure A-28: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model B1.

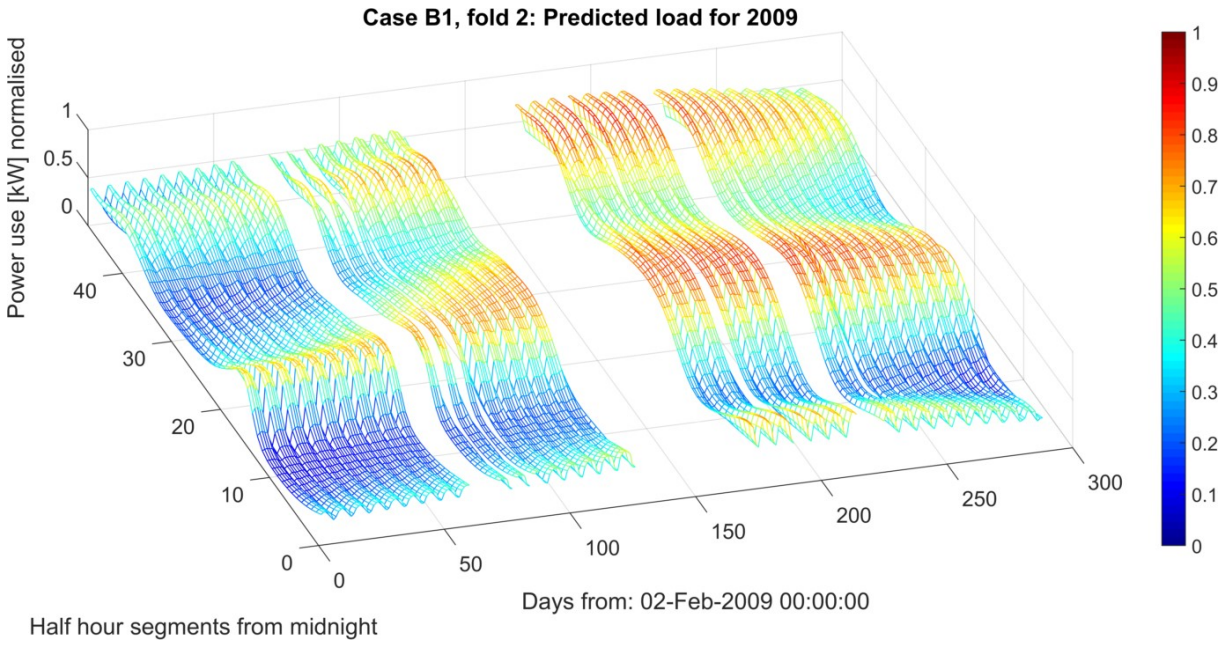

Figure A-29: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model B1.

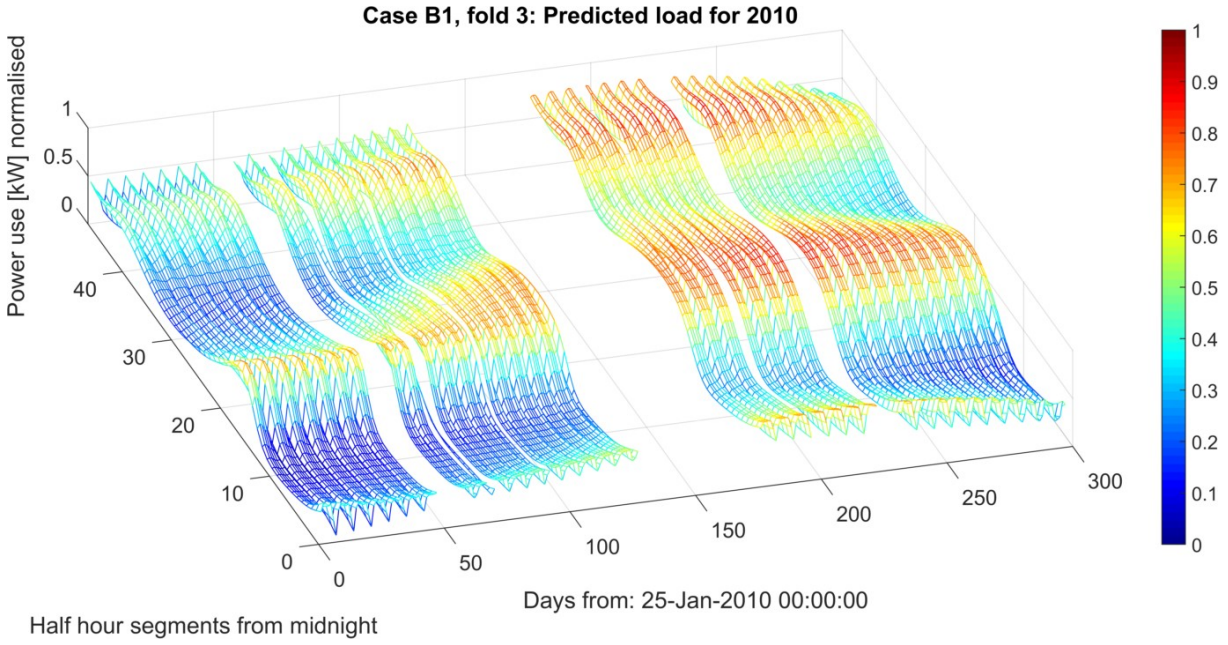

Figure A-30: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model B1.

Appendix A.6 Model B2

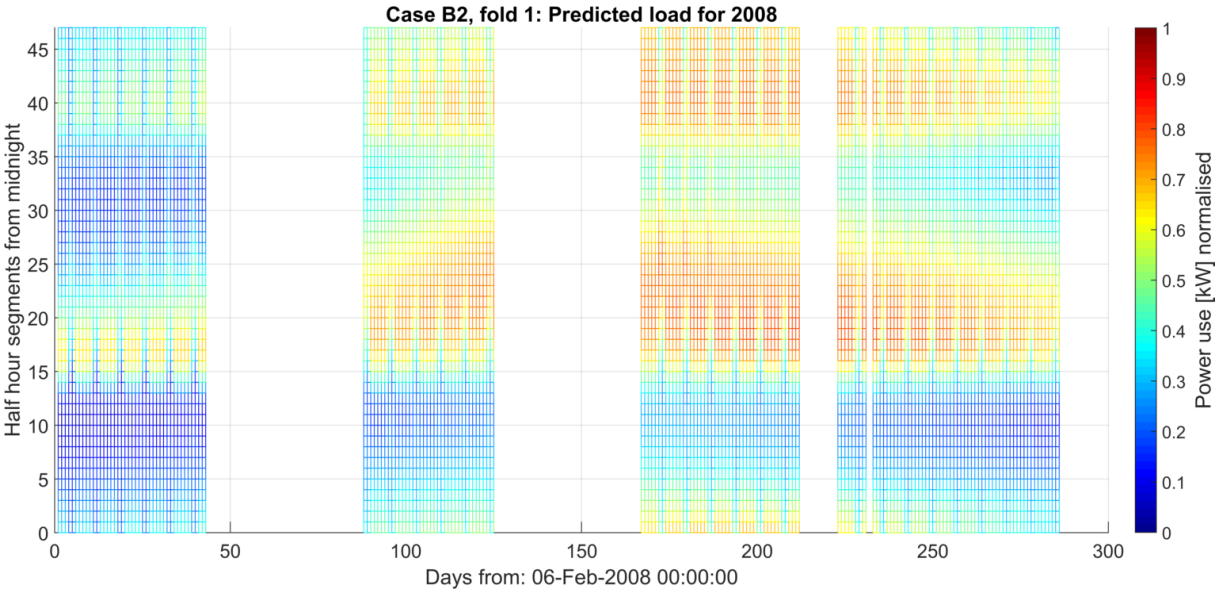

Figure A-31: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model B2.

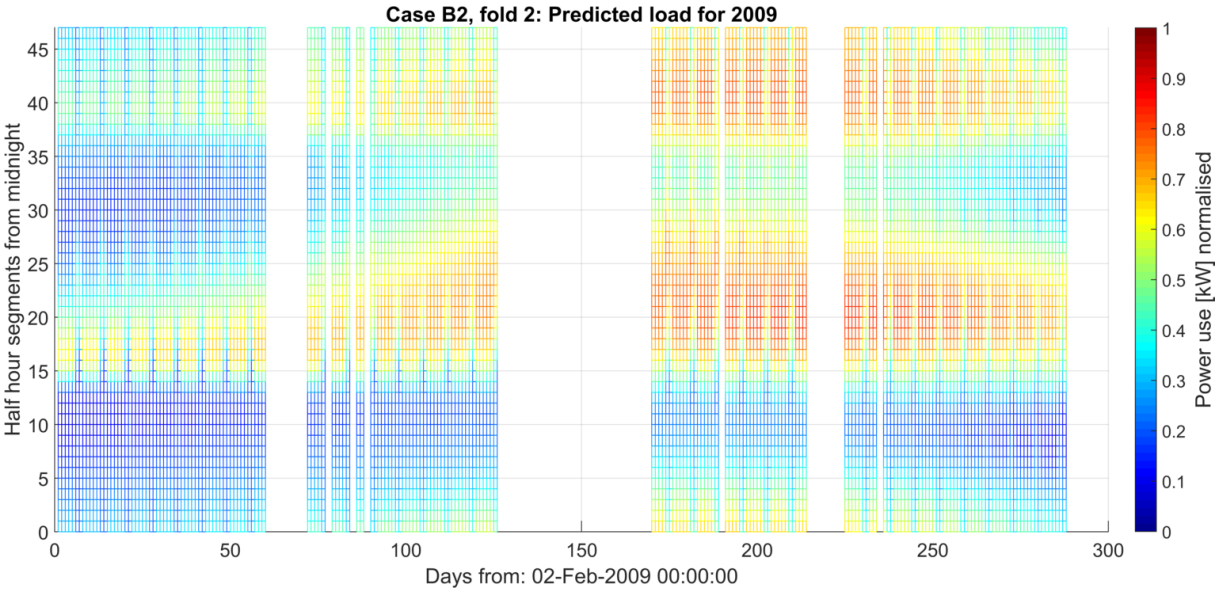

Figure A-32: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model B2.

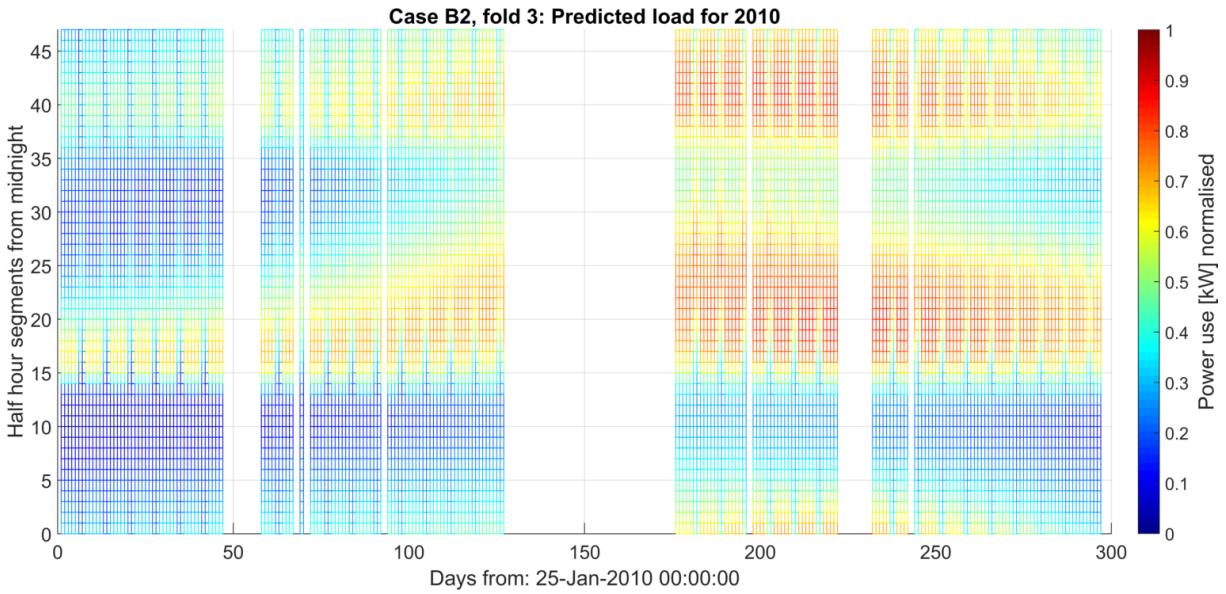

Figure A-33: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model B2.

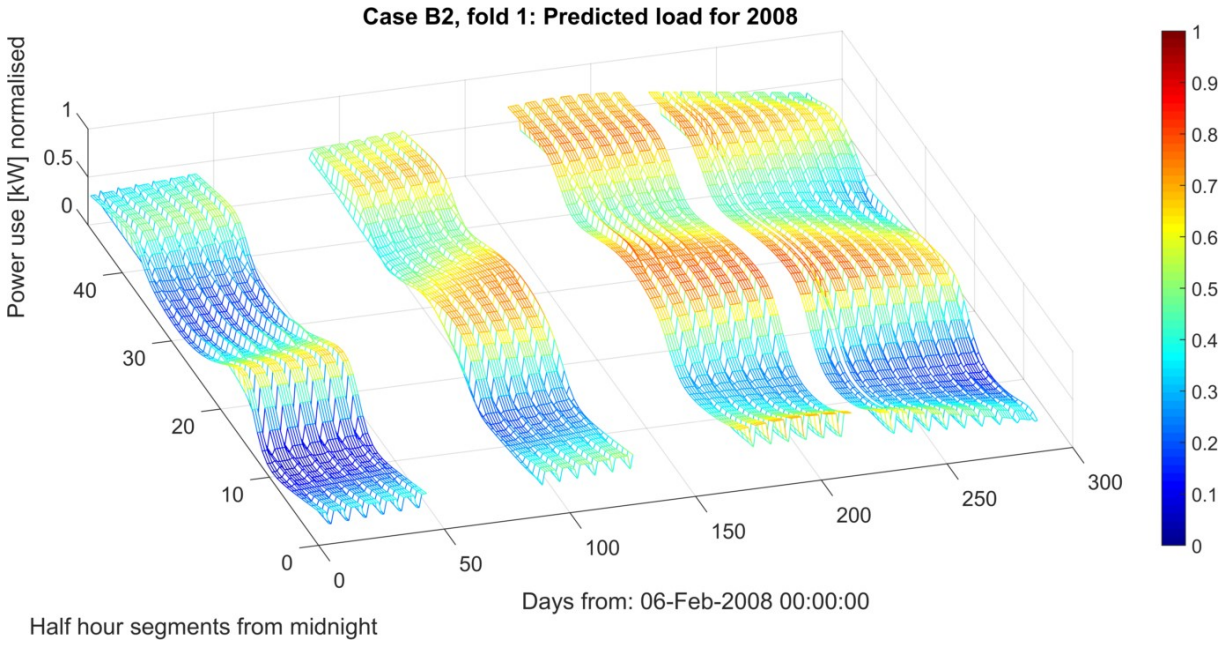

Figure A-34: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model B2.

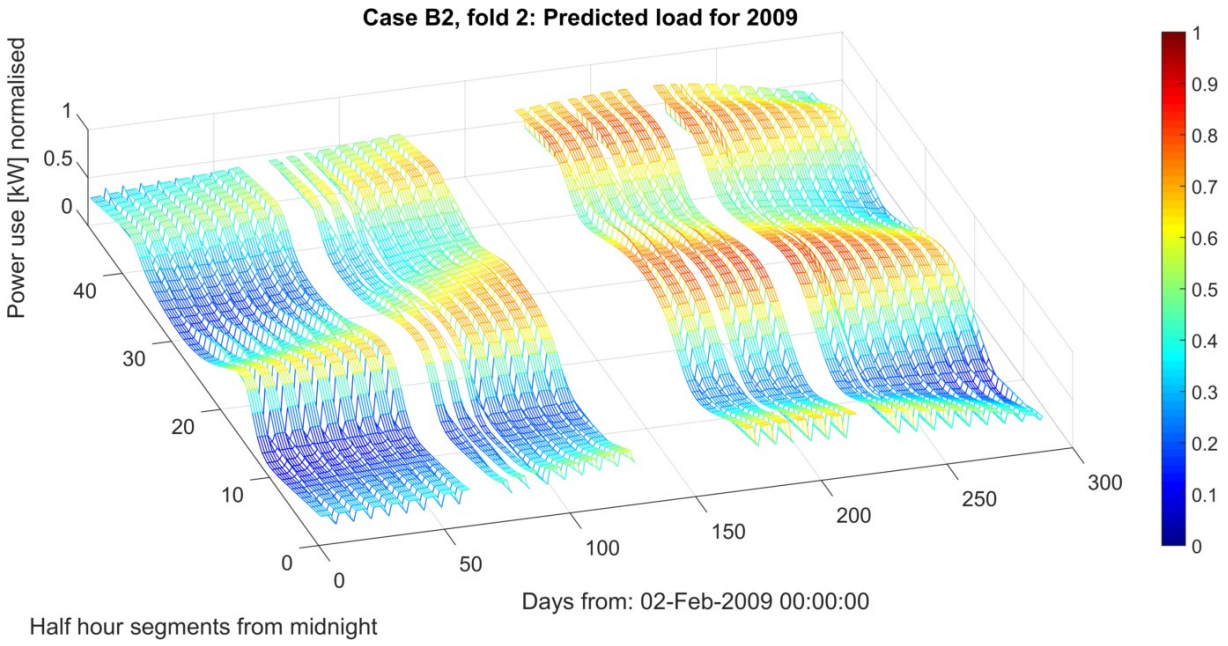

Figure A-35: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model B2.

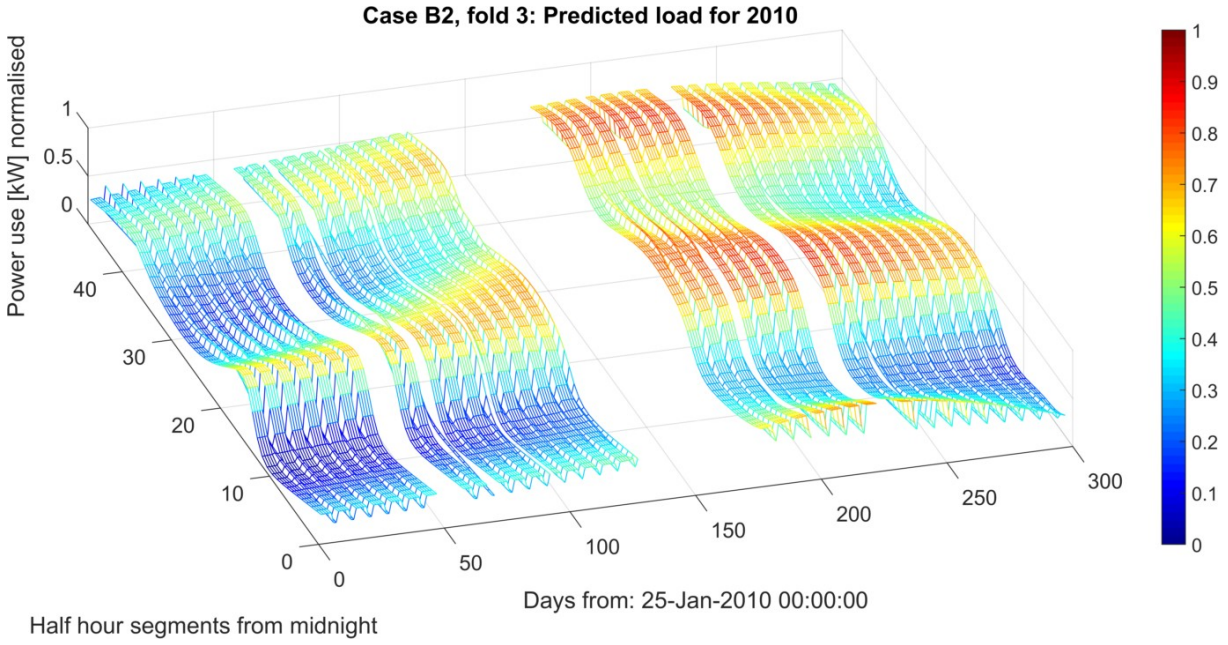

Figure A-36: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model B2.

Appendix A.7 Model B3

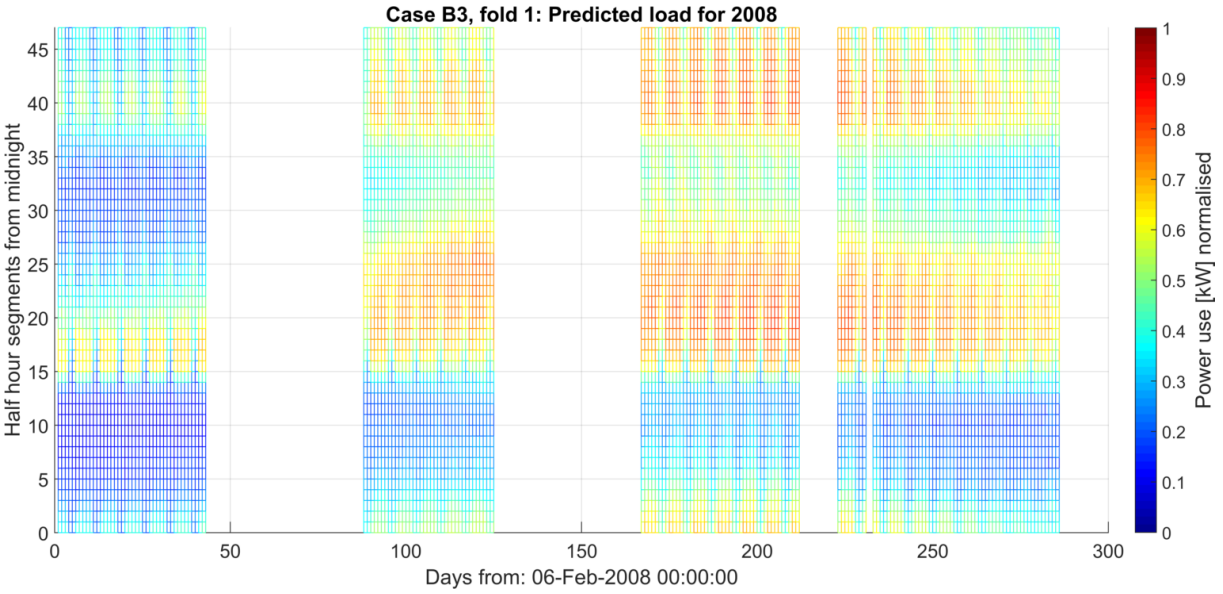

Figure A-37: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model B3.

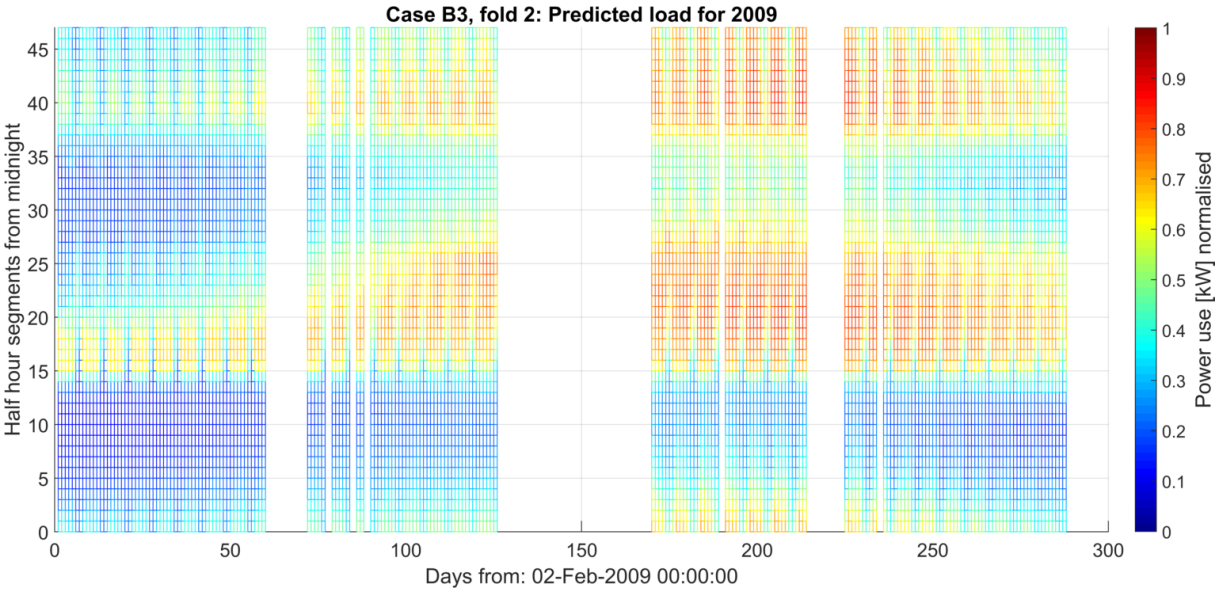

Figure A-38: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model B3.

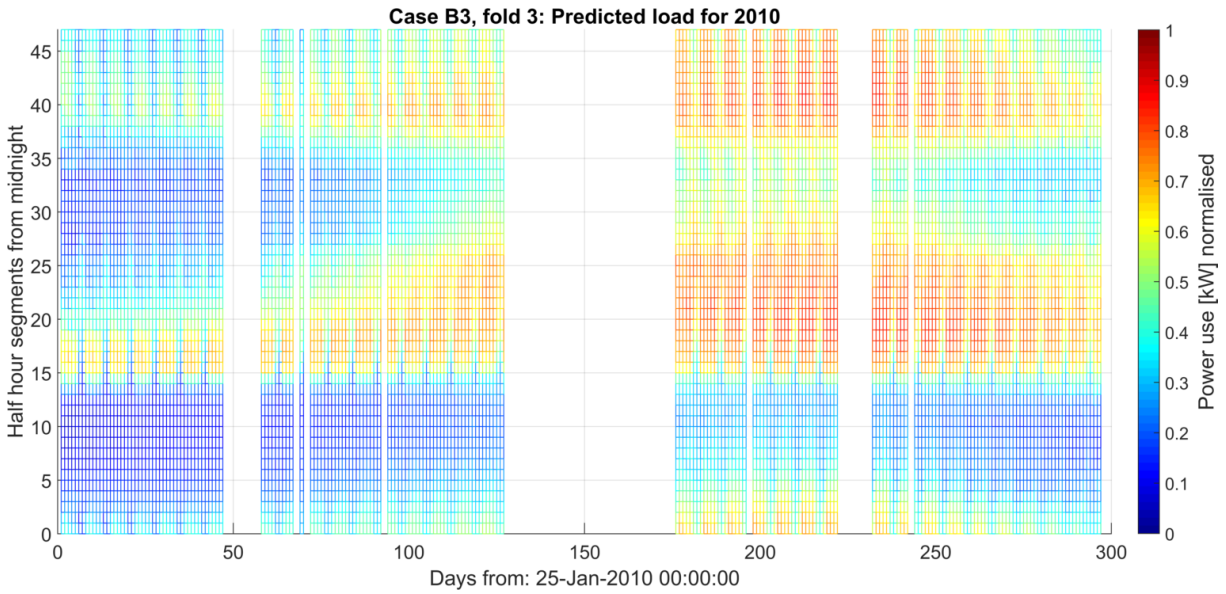

Figure A-39: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model B3.

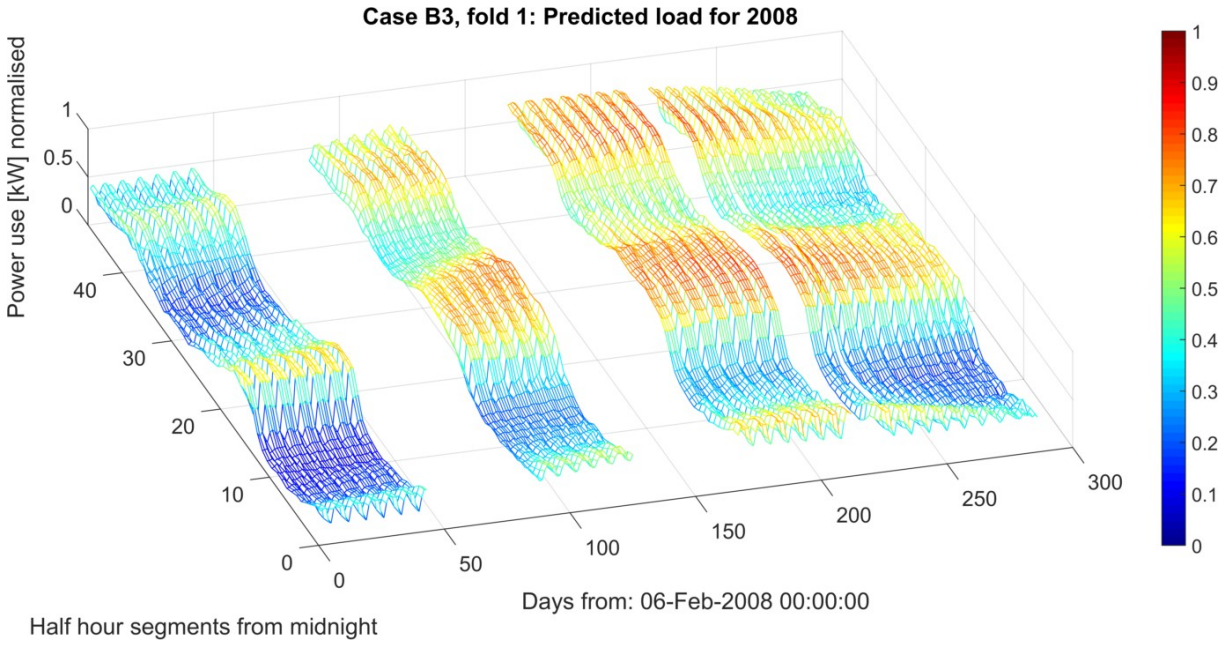

Figure A-40: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model B3.

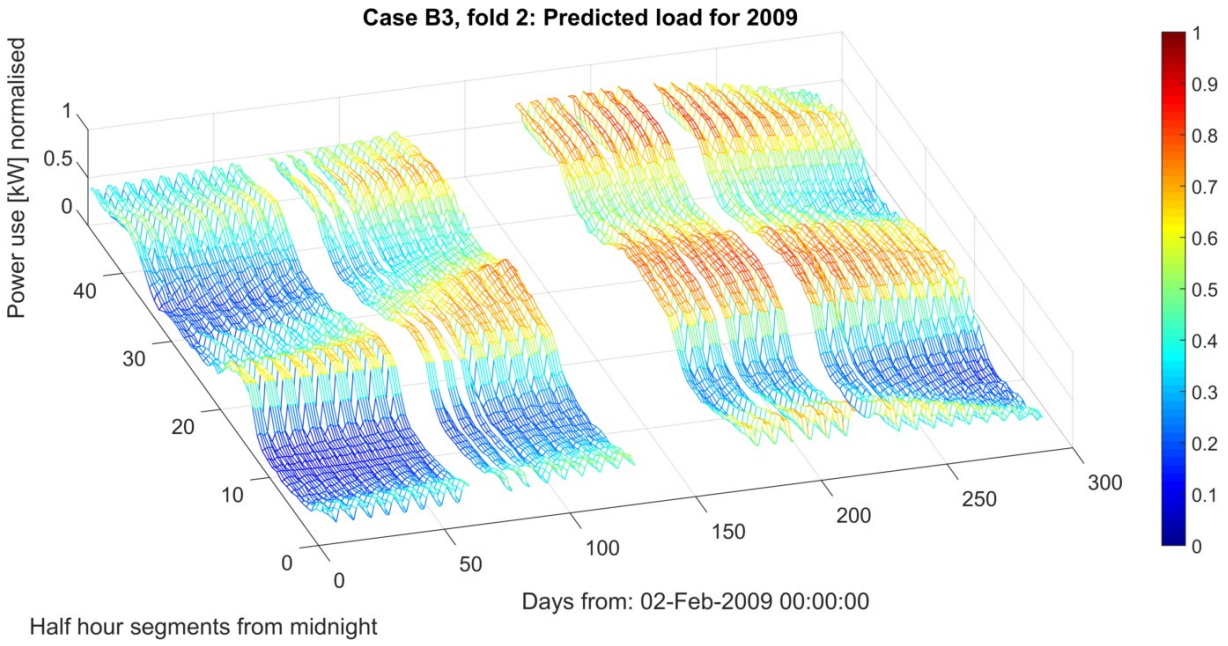

Figure A-41: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model B3.

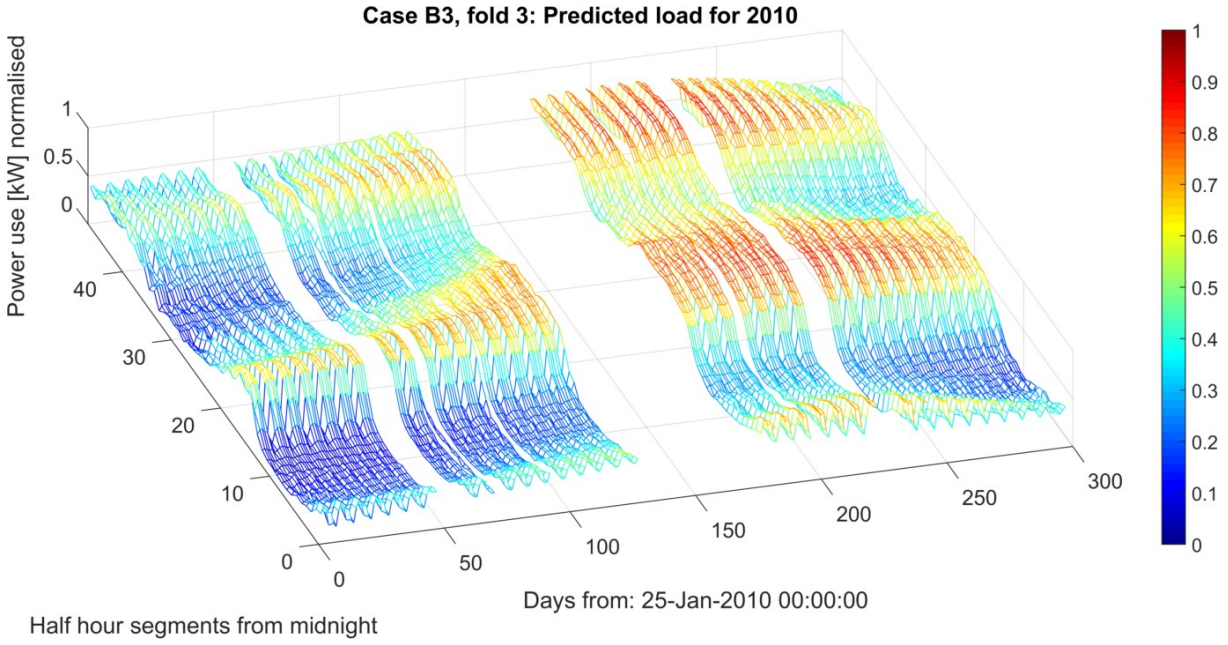

Figure A-42: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model B3.

Appendix A.8 Model C1

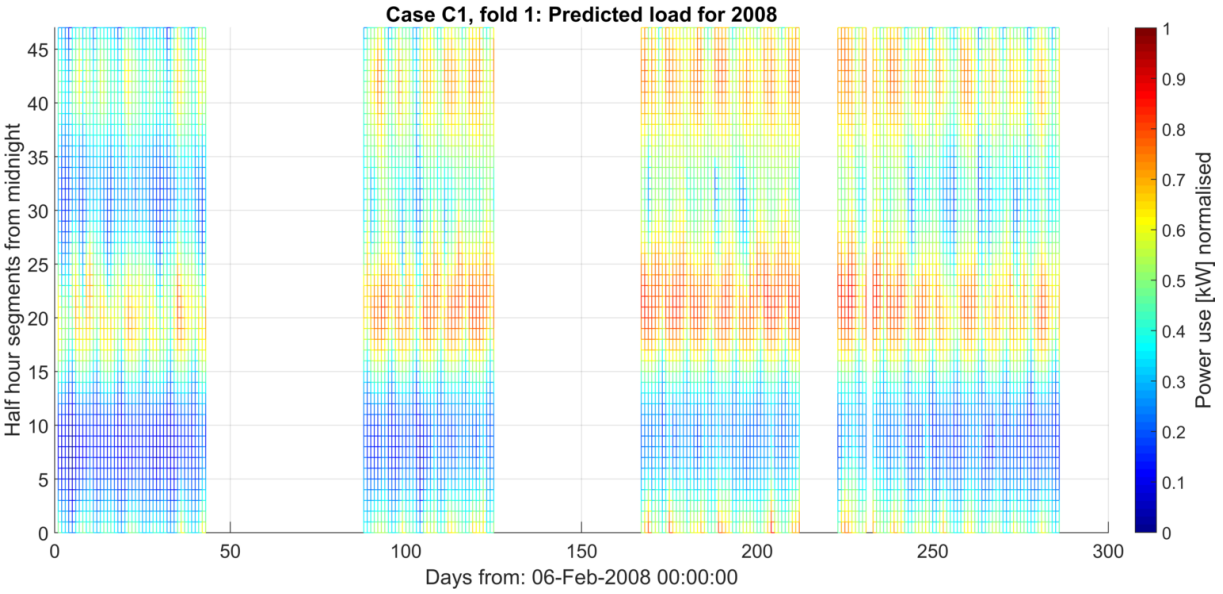

Figure A-43: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model C1.

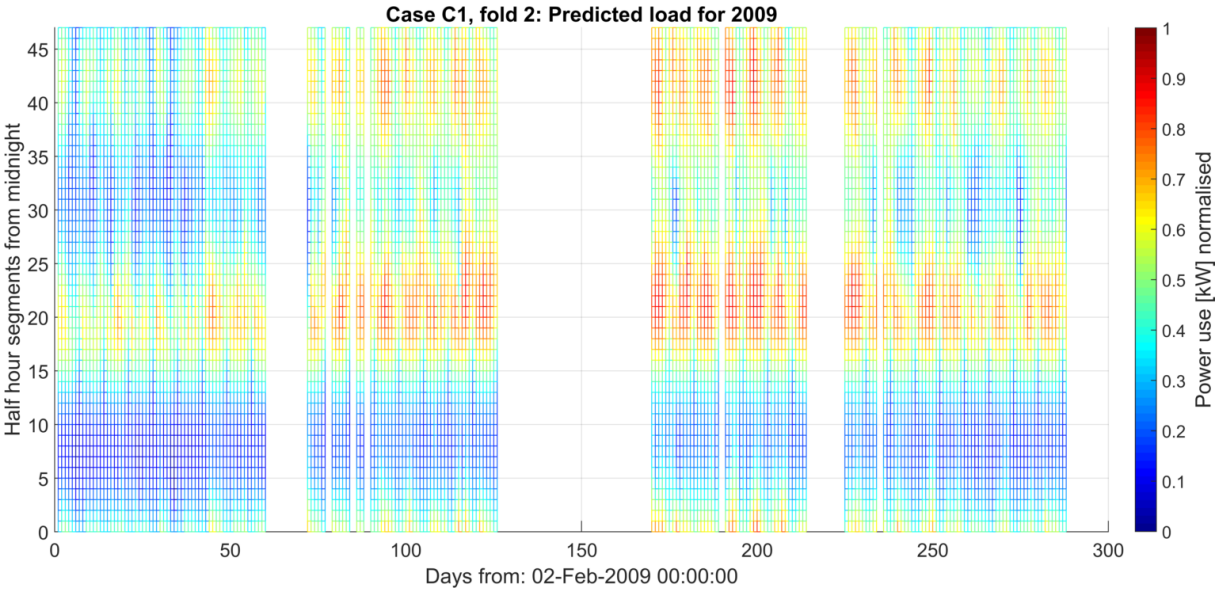

Figure A-44: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model C1.

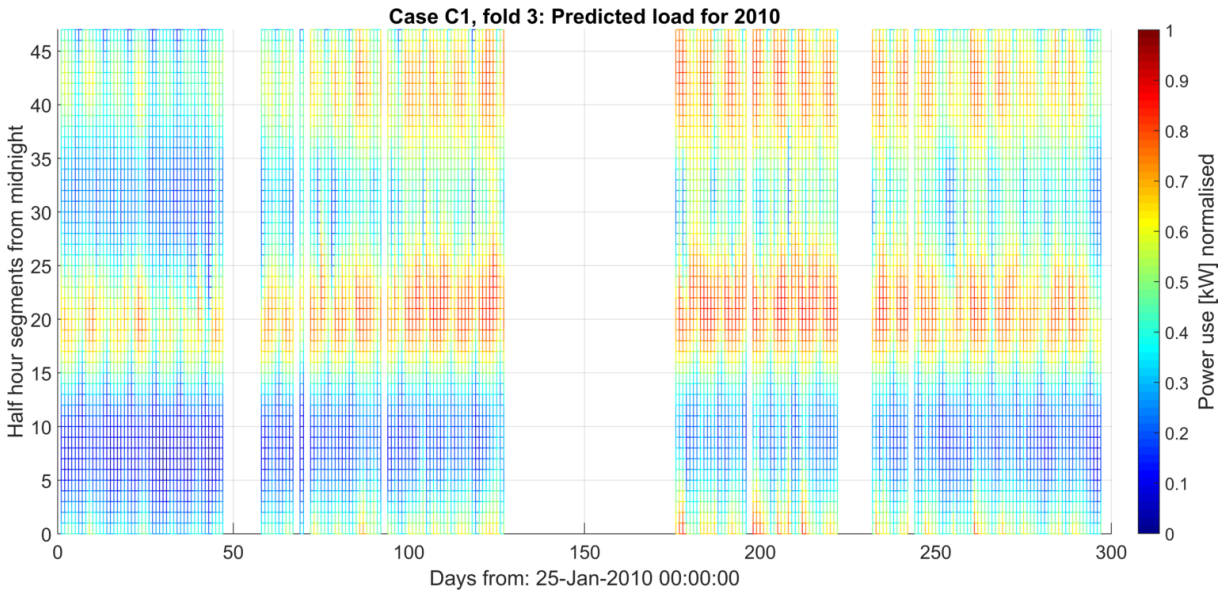

Figure A-45: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model C1.

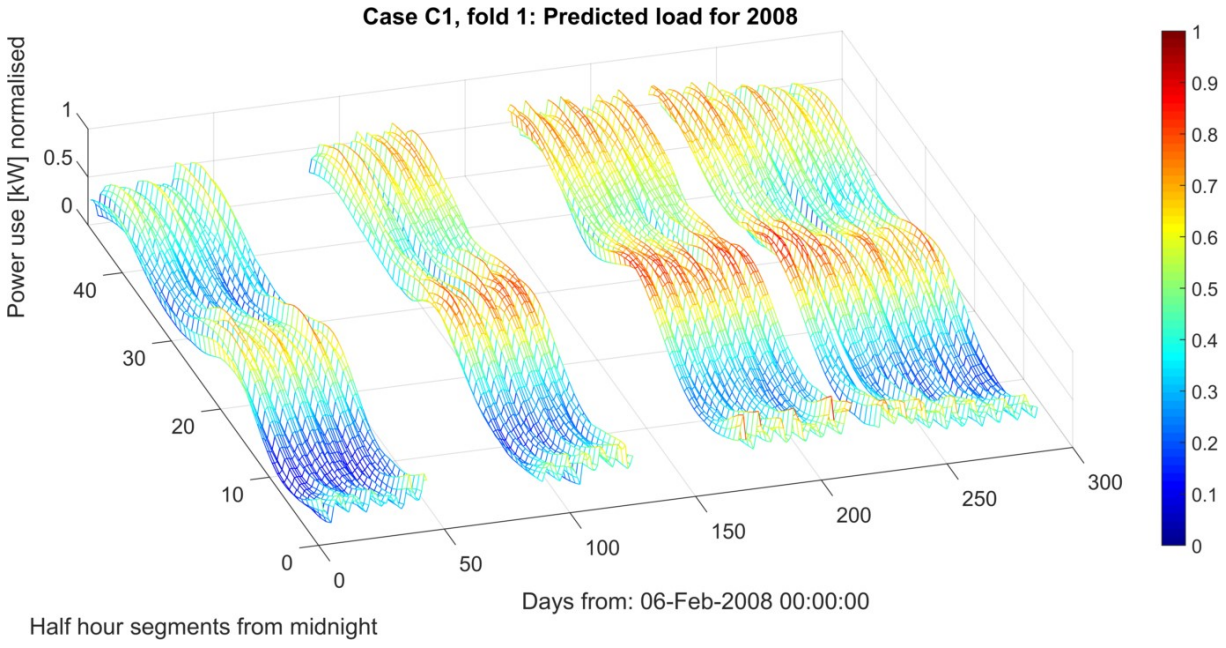

Figure A-46: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model C1.

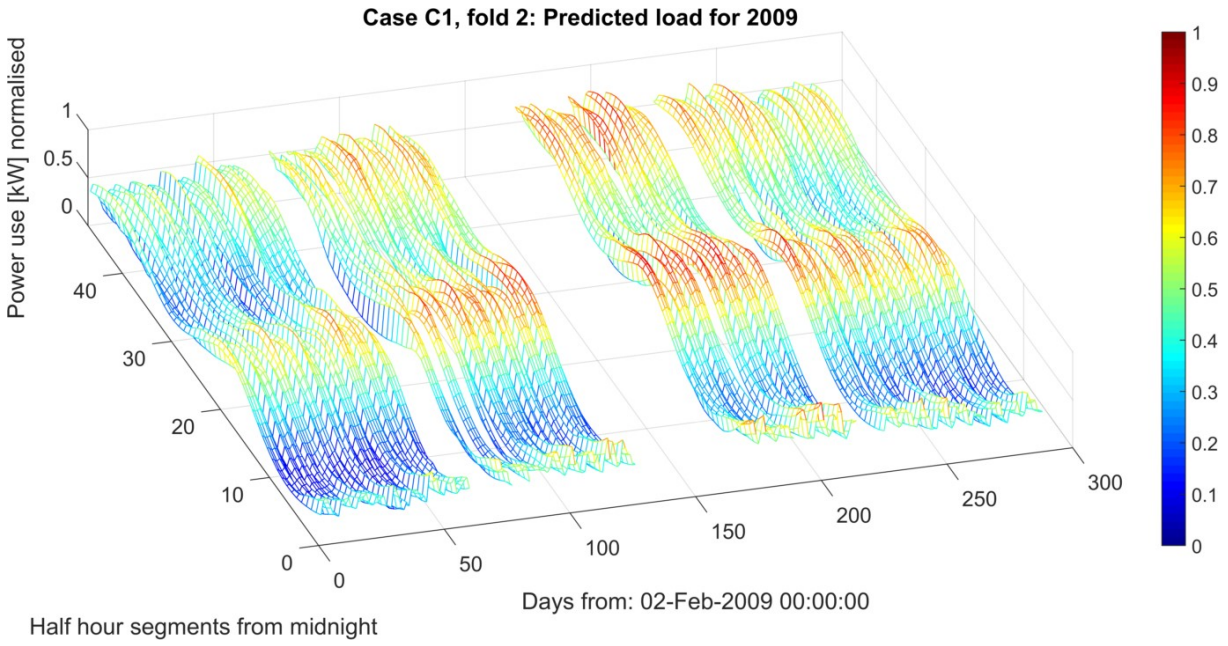

Figure A-47: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model C1.

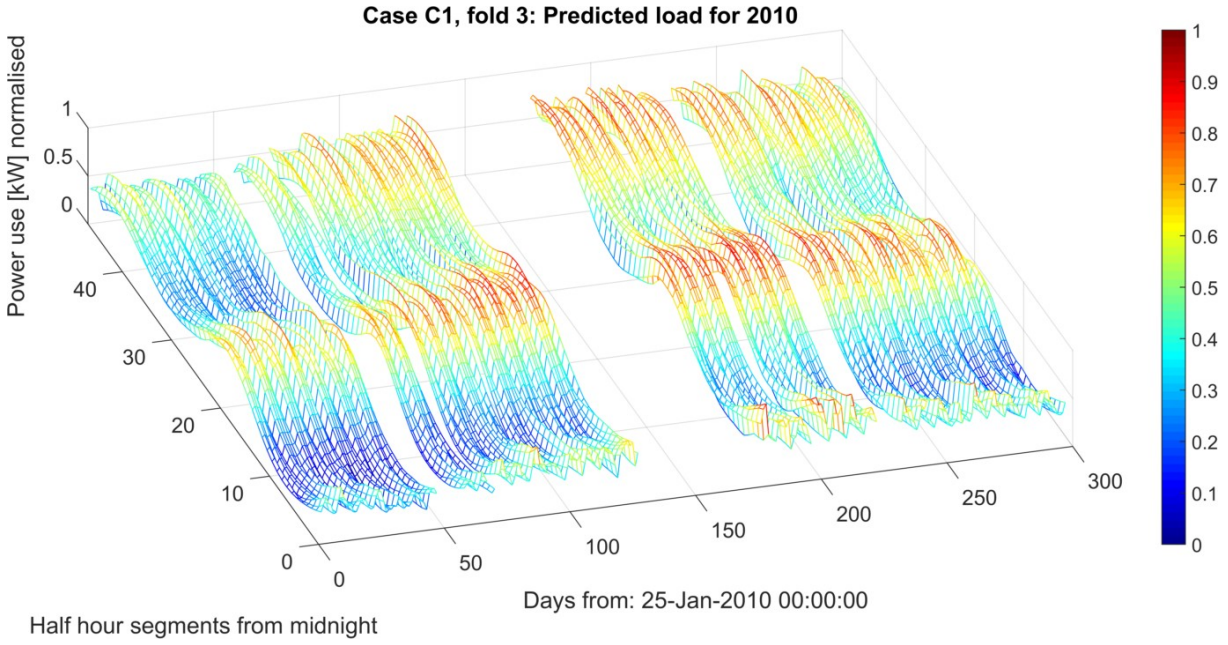

Figure A-48: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model C1.

Appendix A.9 Model C2

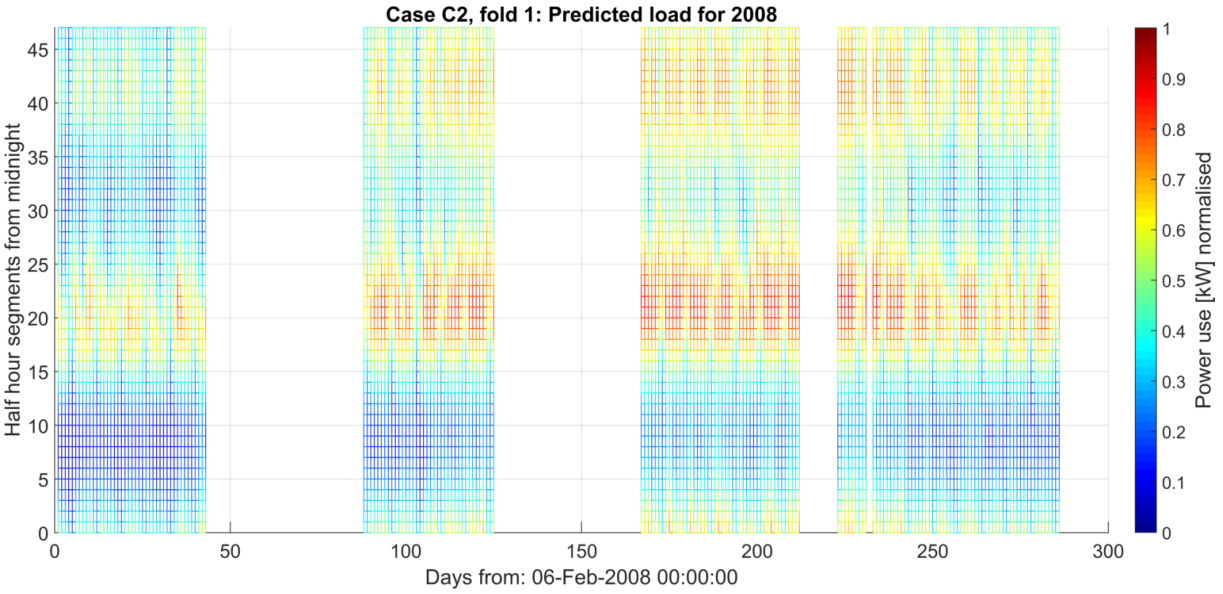

Figure A-49: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model C2.

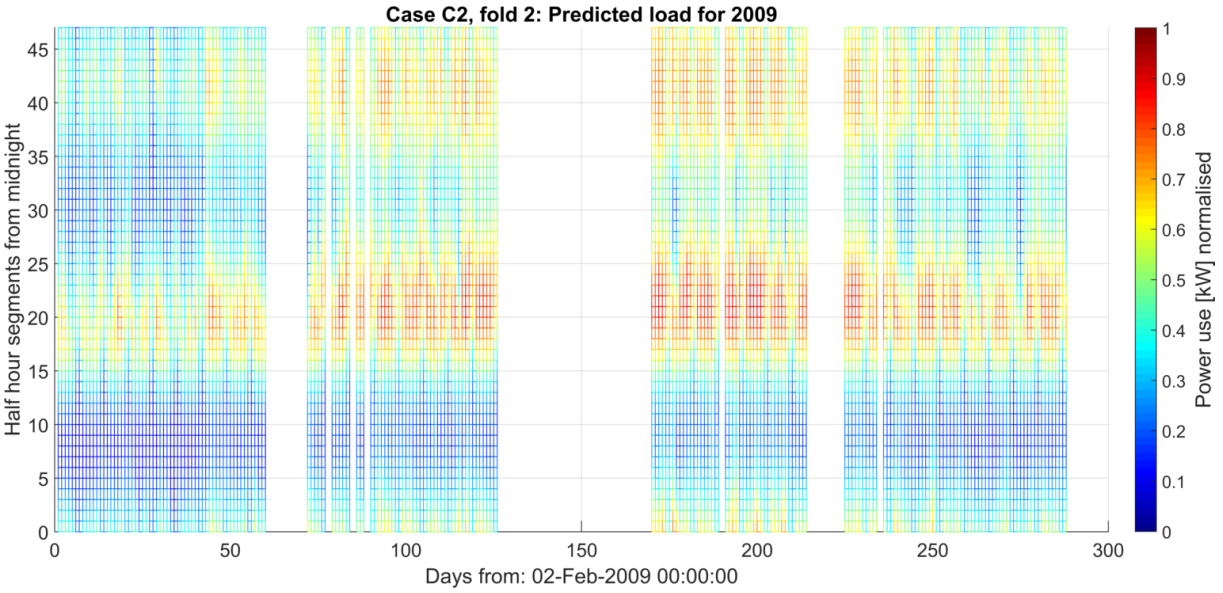

Figure A-50: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model C2.

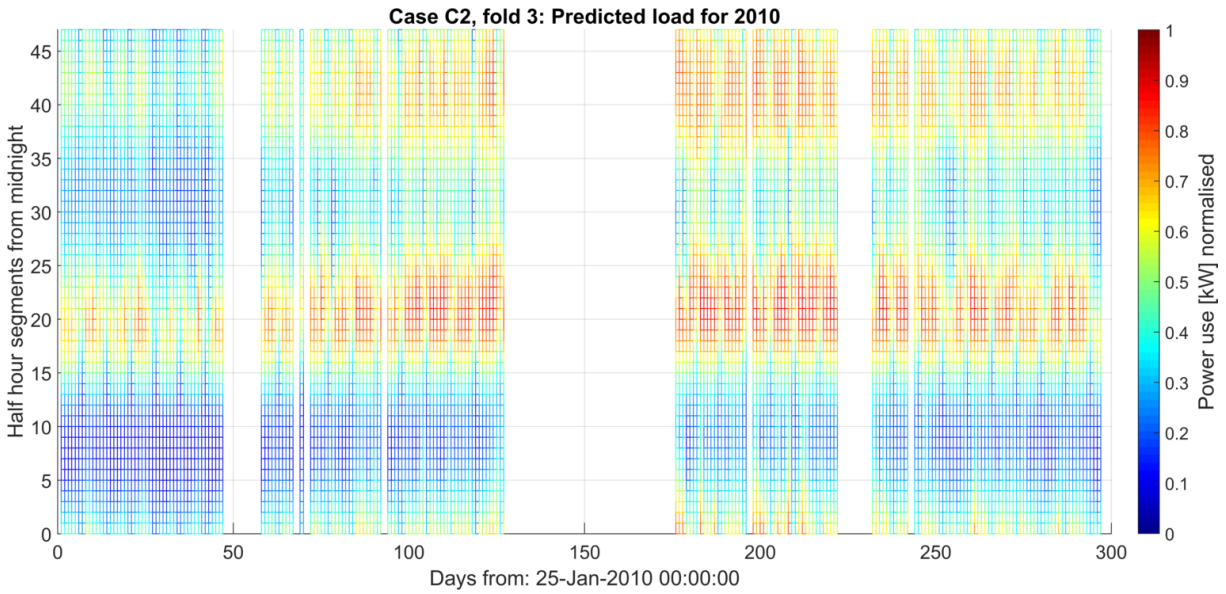

Figure A-51: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model C2.

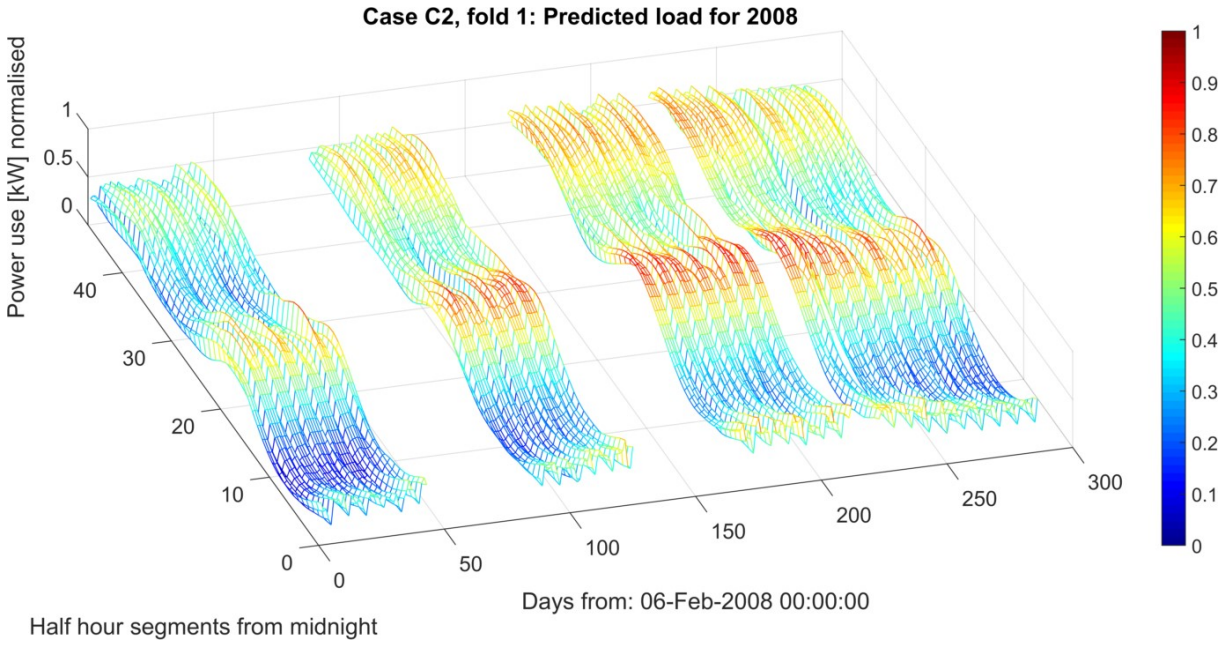

Figure A-52: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model C2.

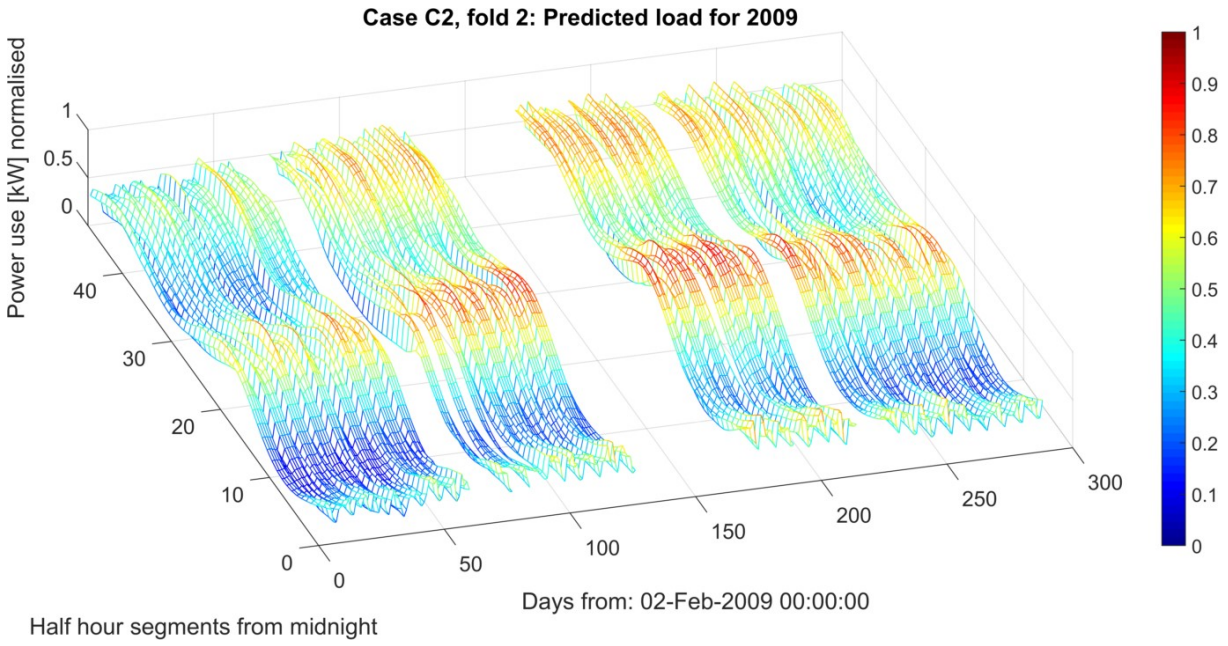

Figure A-53: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model C2.

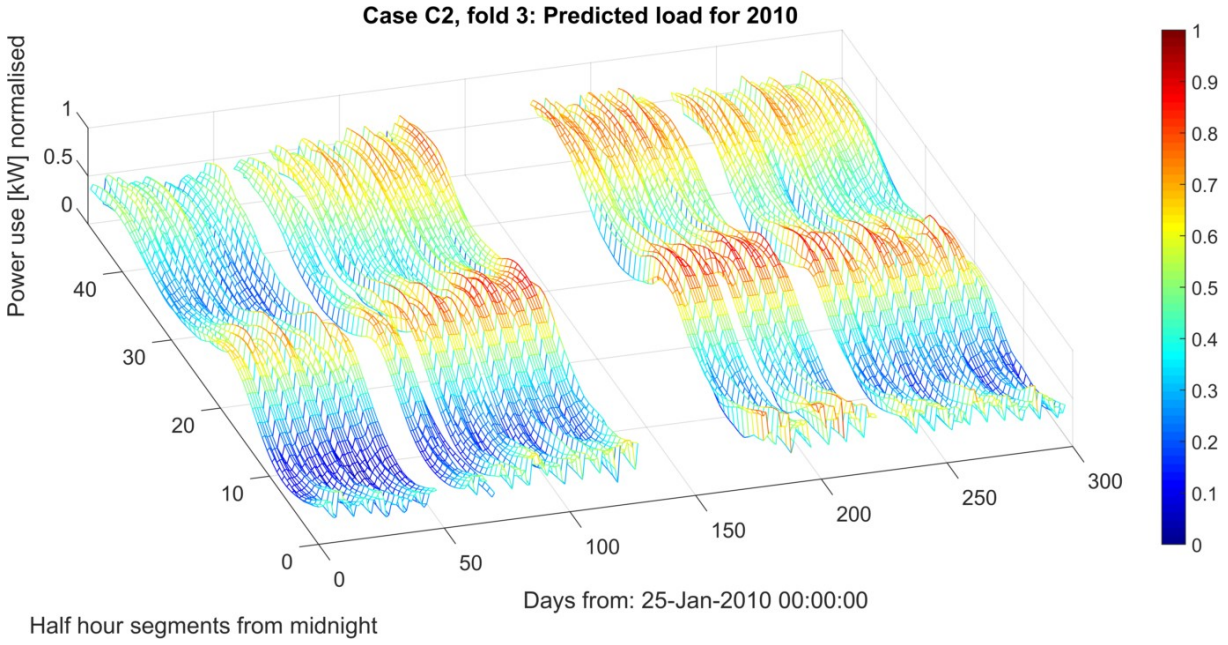

Figure A-54: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model C2.

Appendix A.10 Model C3

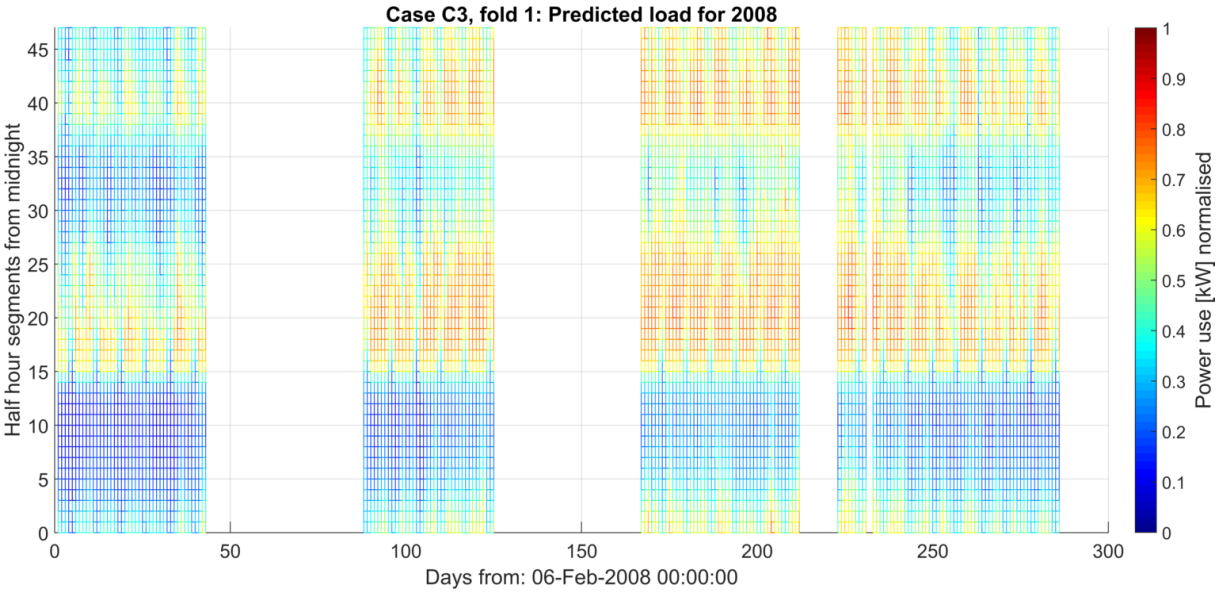

Figure A-55: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model C3.

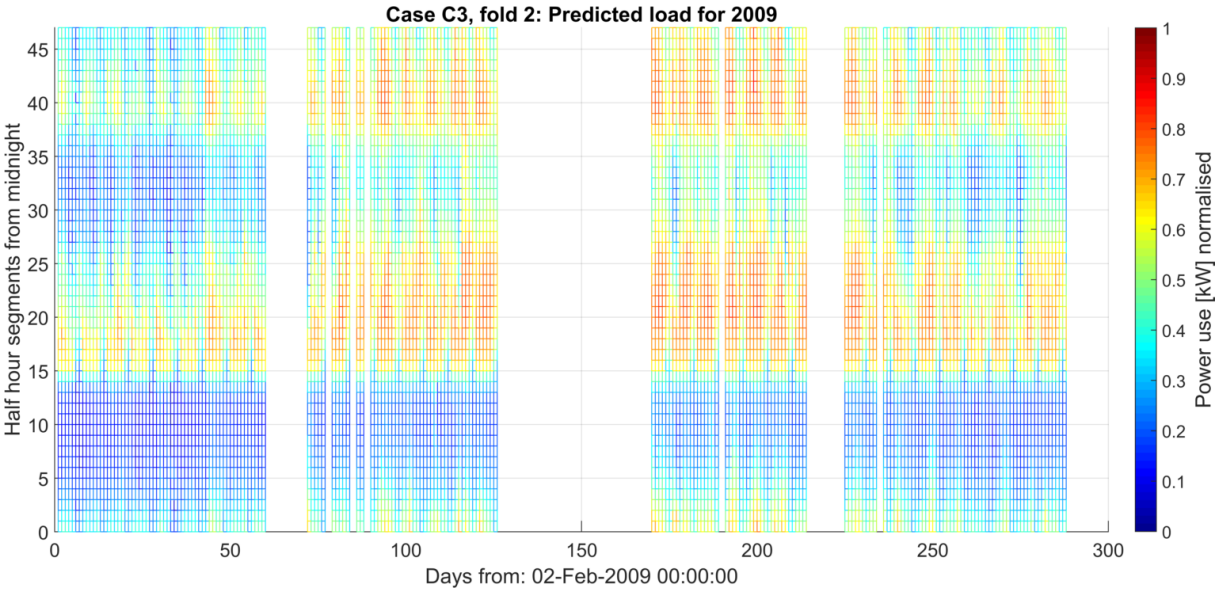

Figure A-56: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model C3.

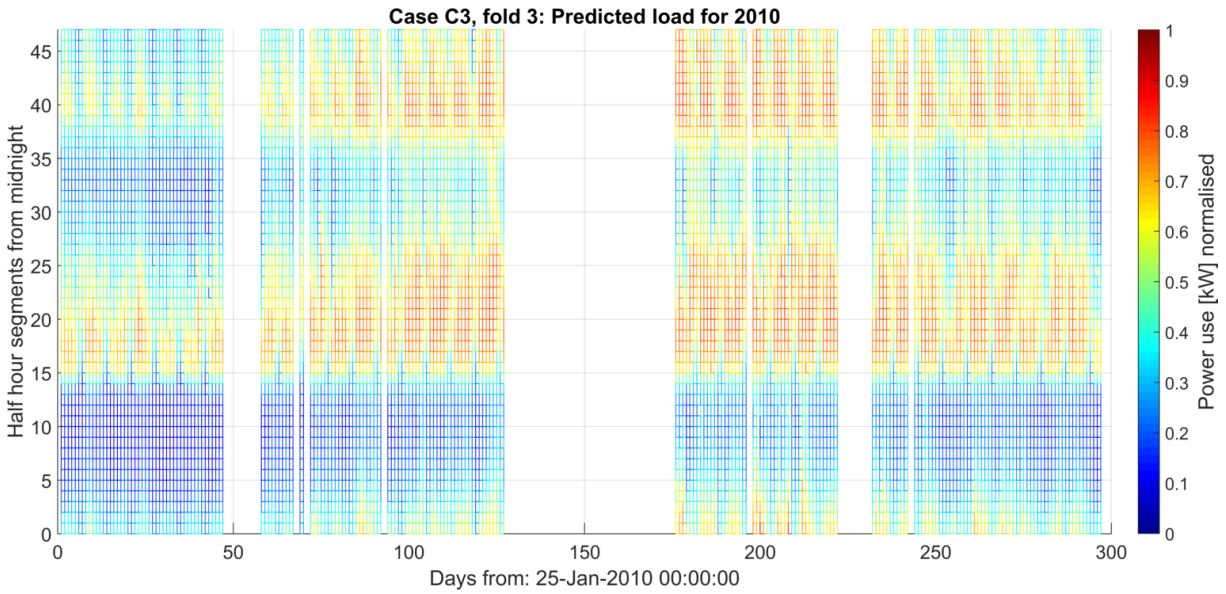

Figure A-57: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model C3.

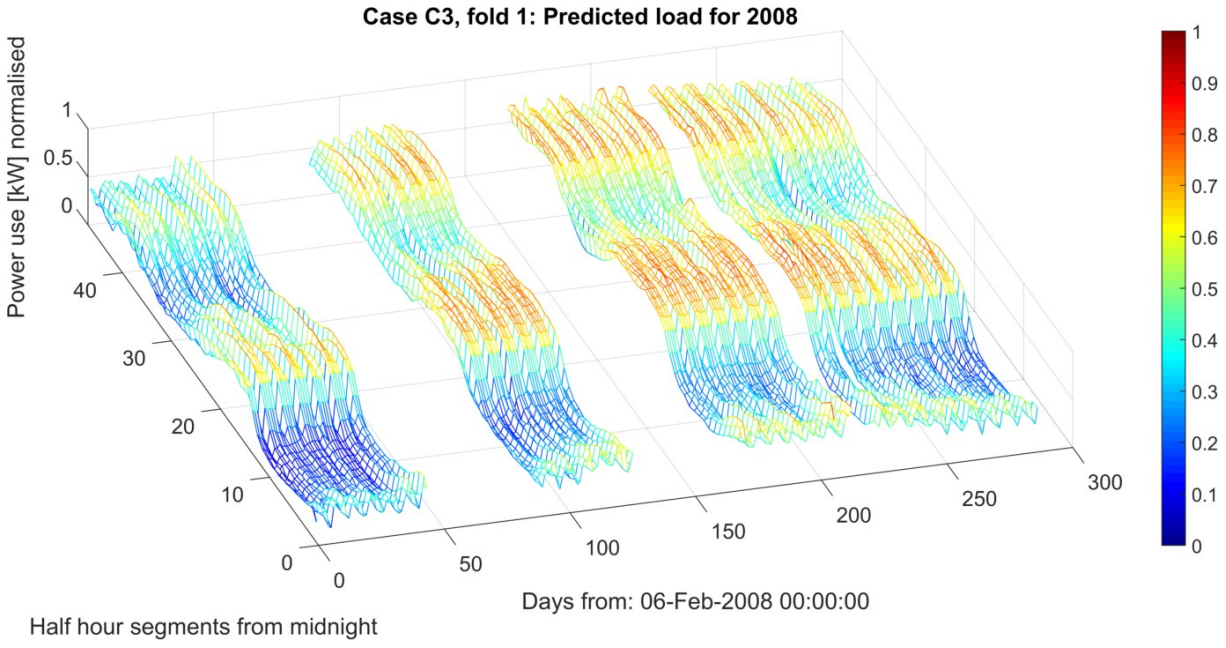

Figure A-58: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model C3.

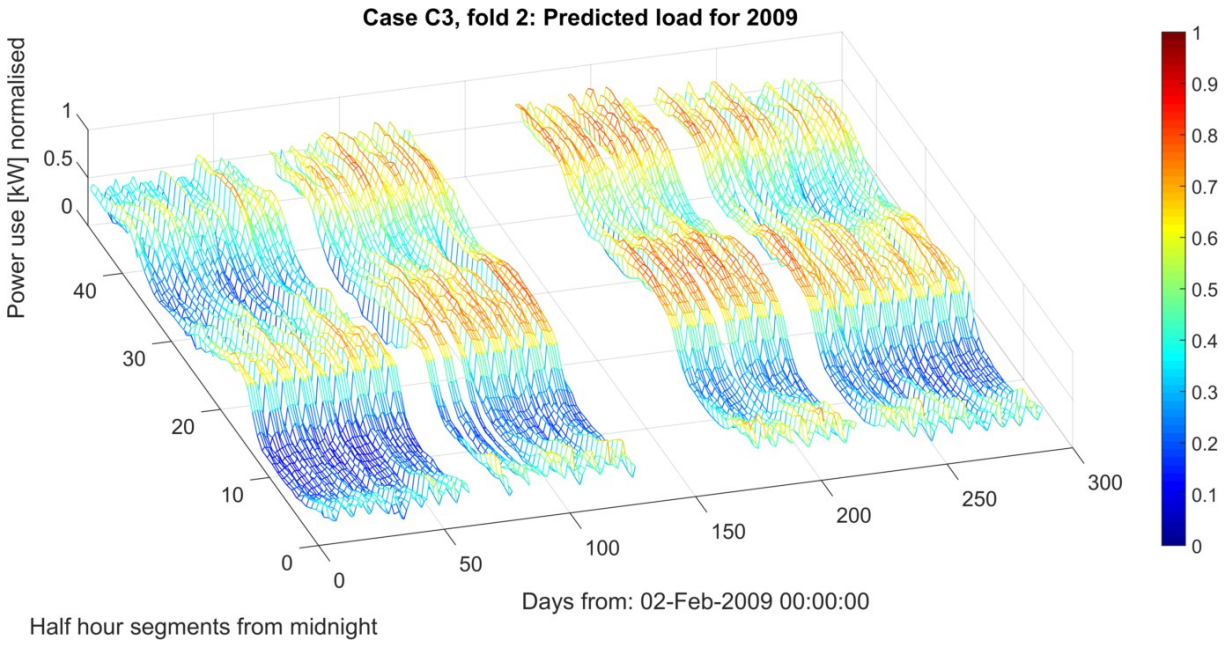

Figure A-59: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model C3.

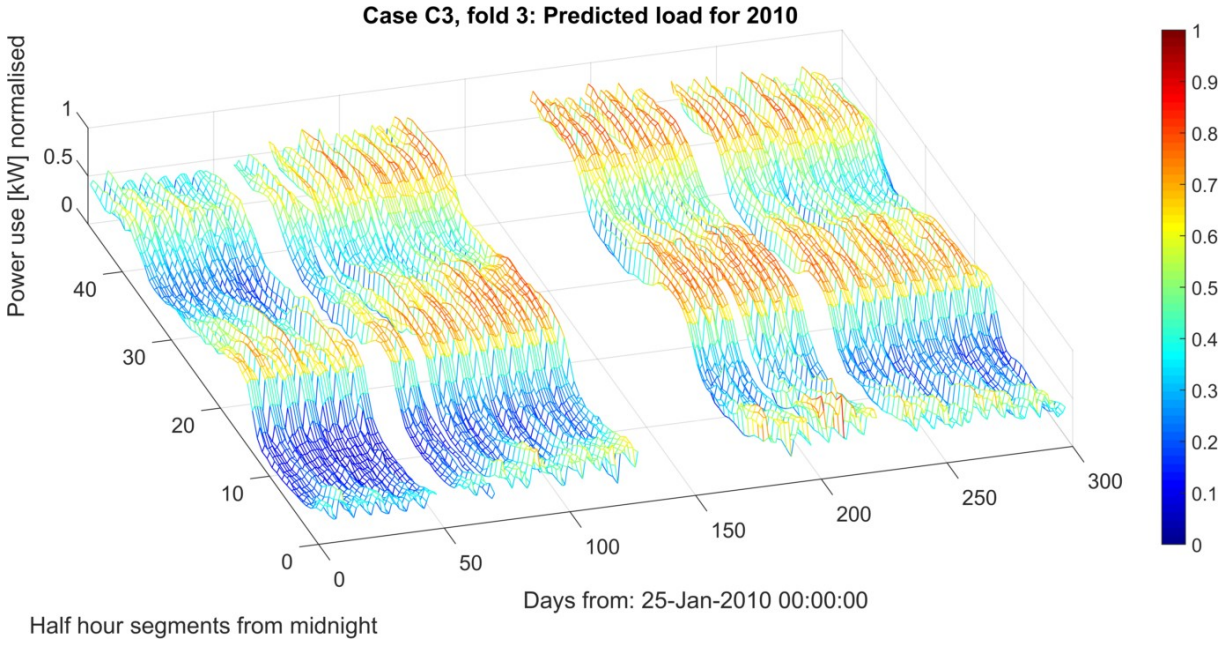

Figure A-60: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model C3.

Appendix A.11 Model D1

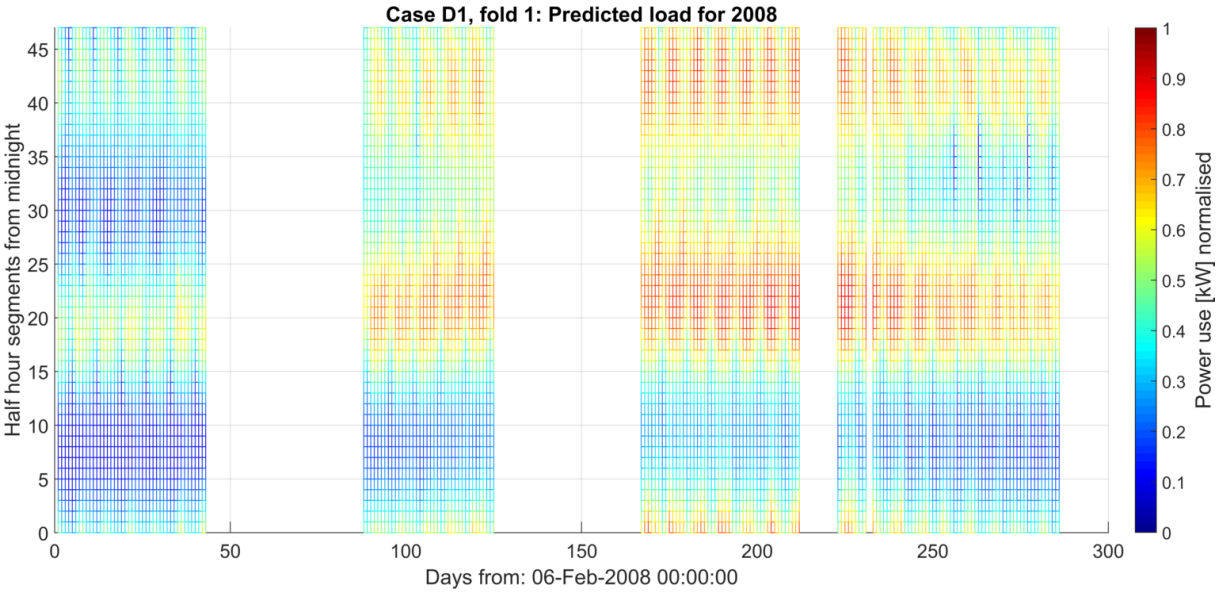

Figure A-61: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model D1.

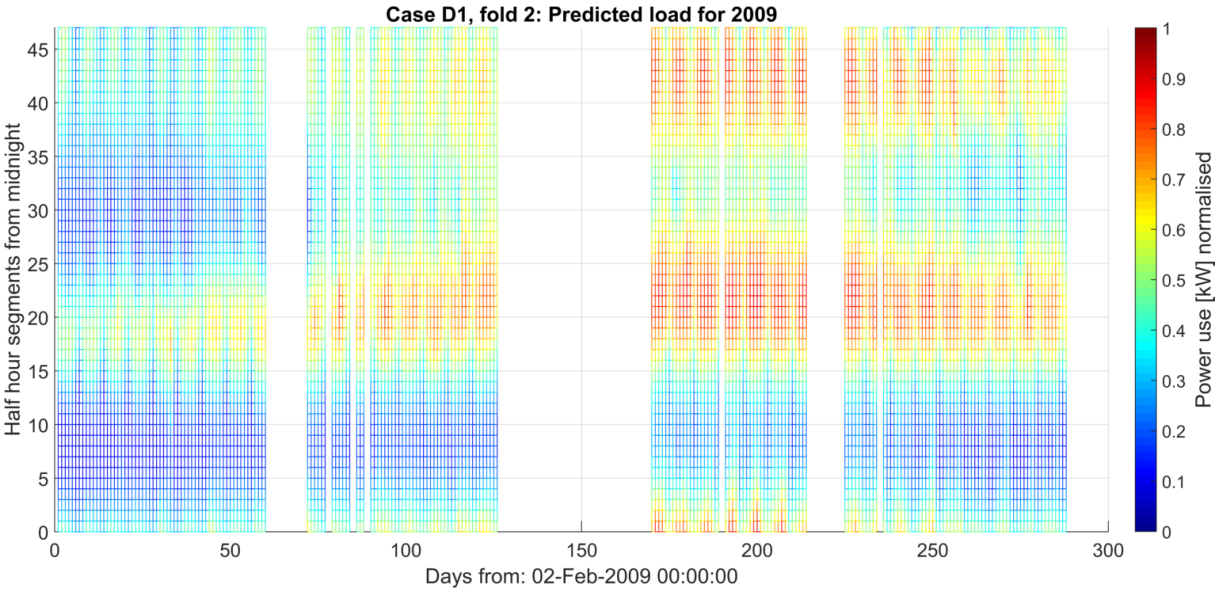

Figure A-62: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model D1.

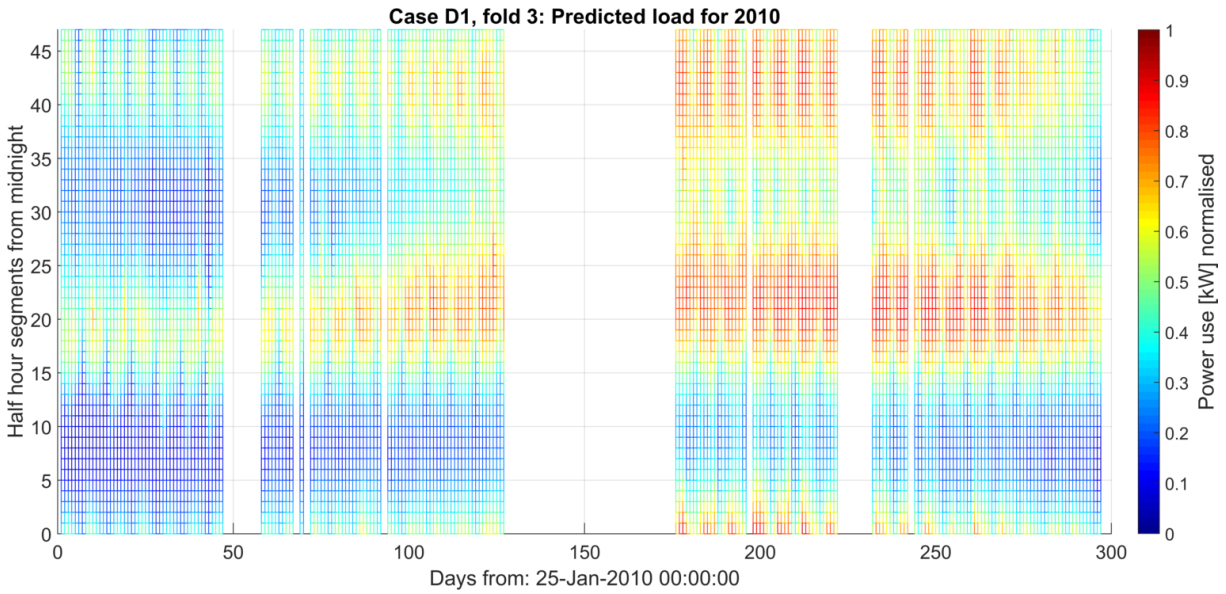

Figure A-63: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model D1.

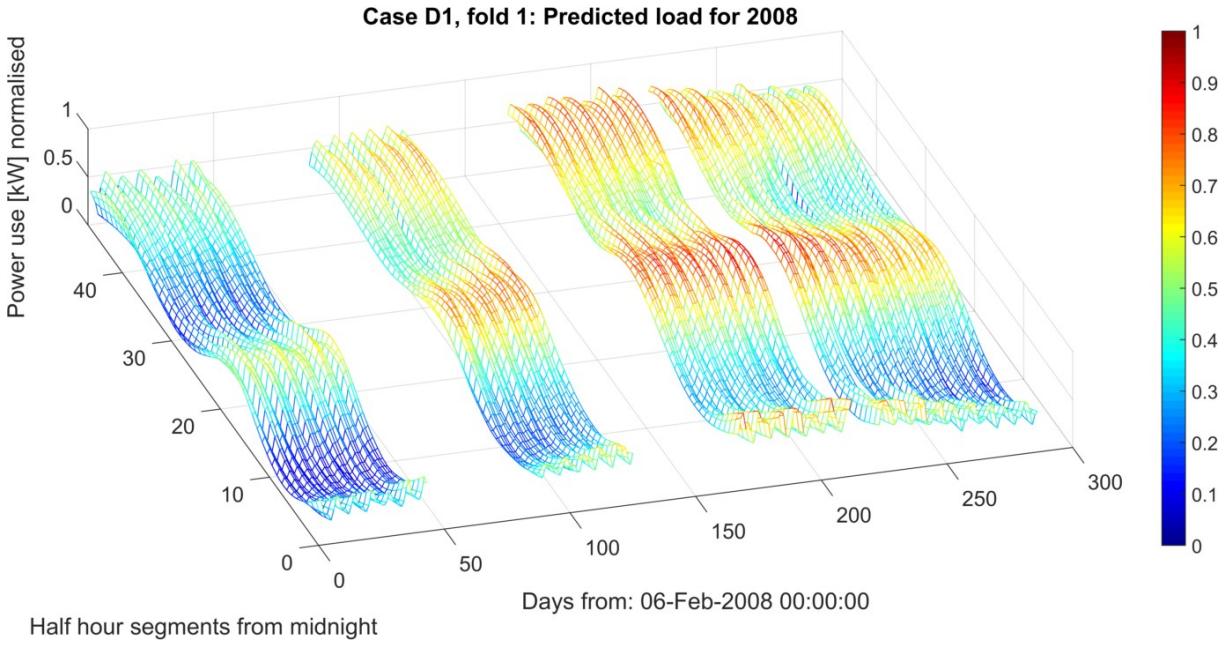

Figure A-64: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model D1.

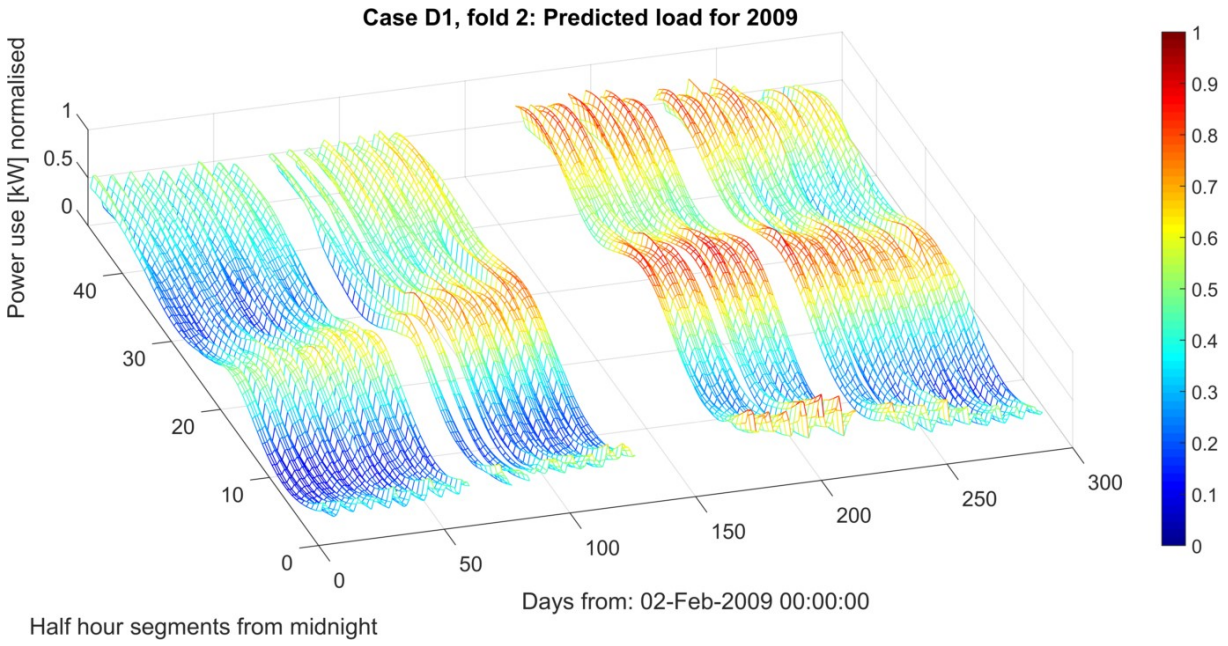

Figure A-65: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model D1.

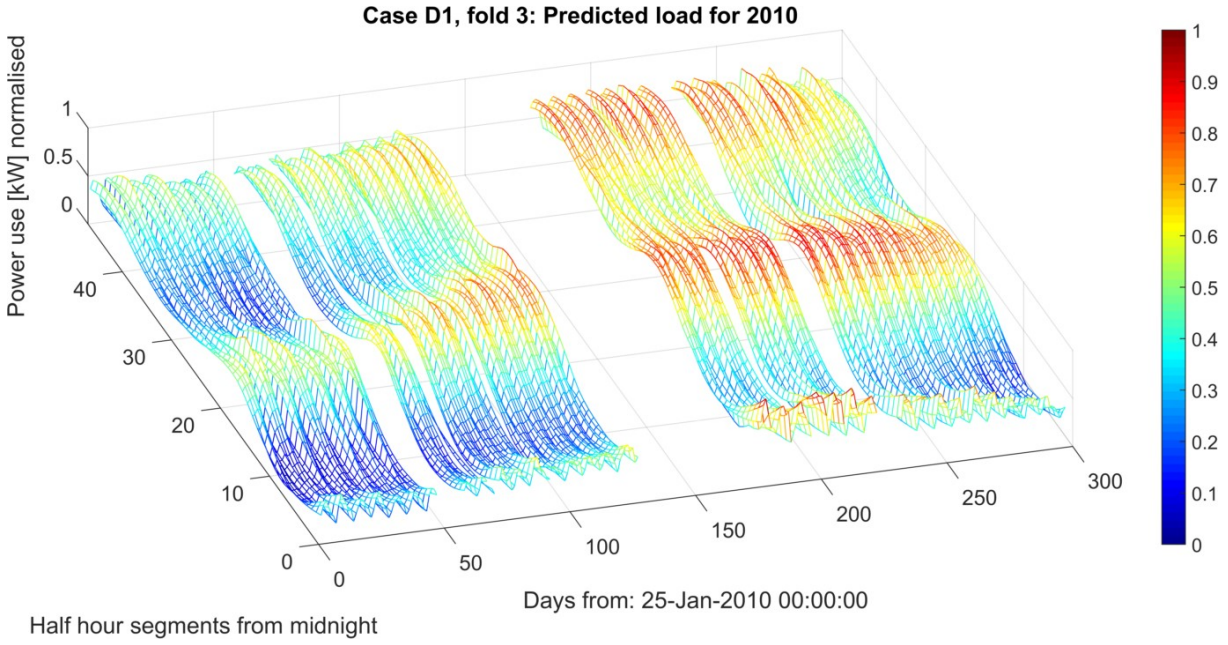

Figure A-66: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model D1.

Appendix A.12 Model D2

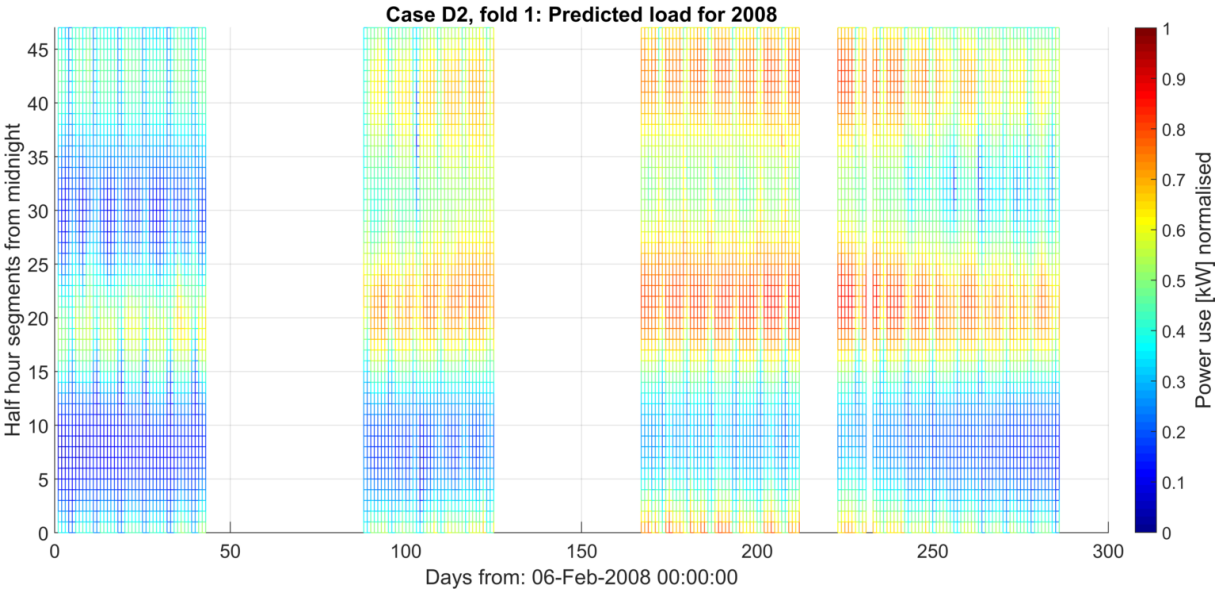

Figure A-67: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model D2.

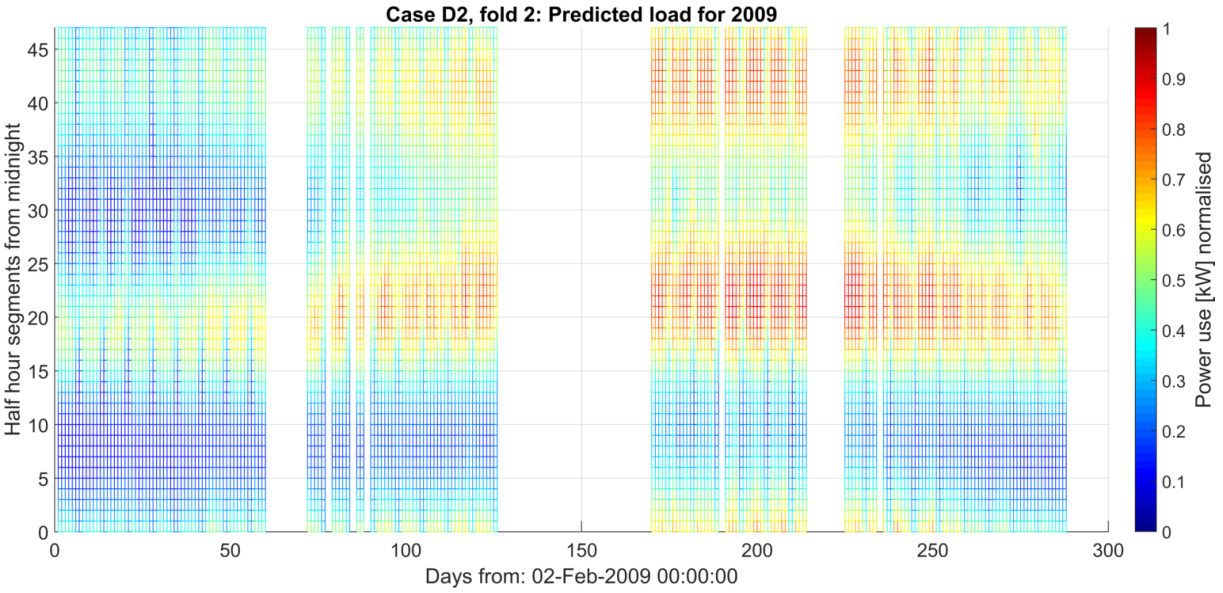

Figure A-68: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model D2.

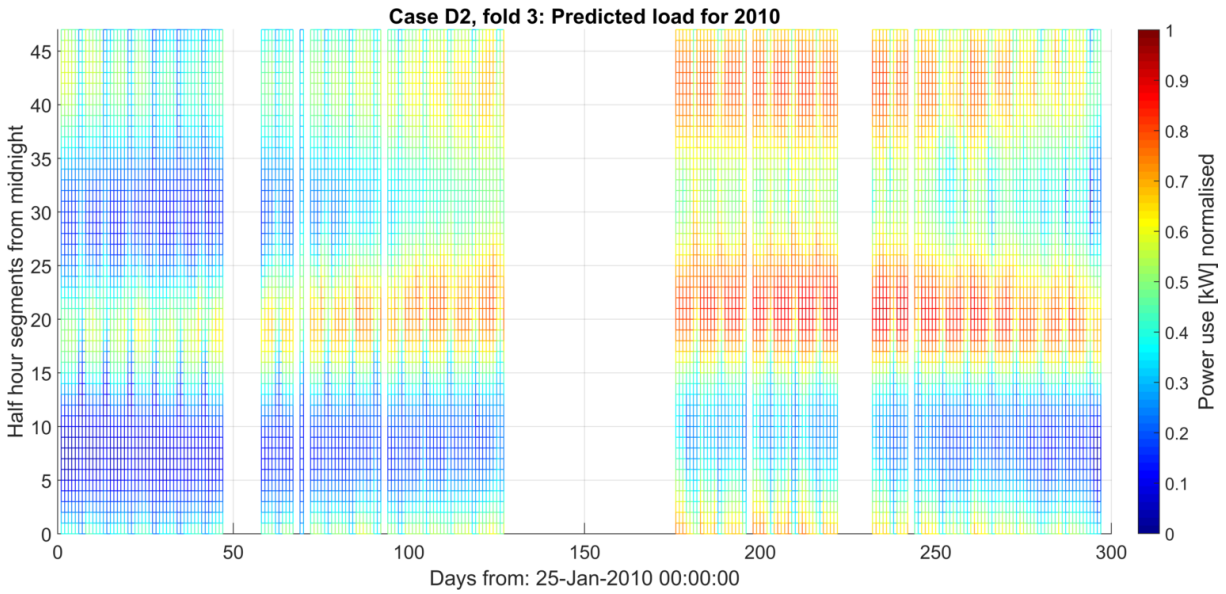

Figure A-69: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model D2.

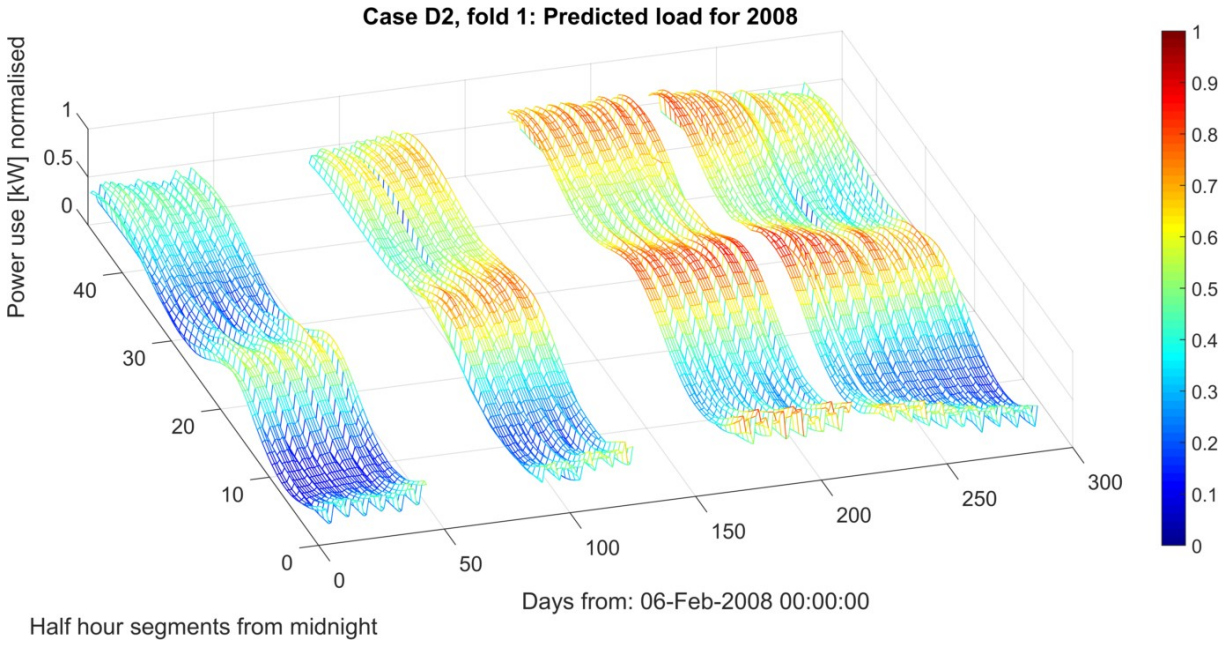

Figure A-70: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model D2.

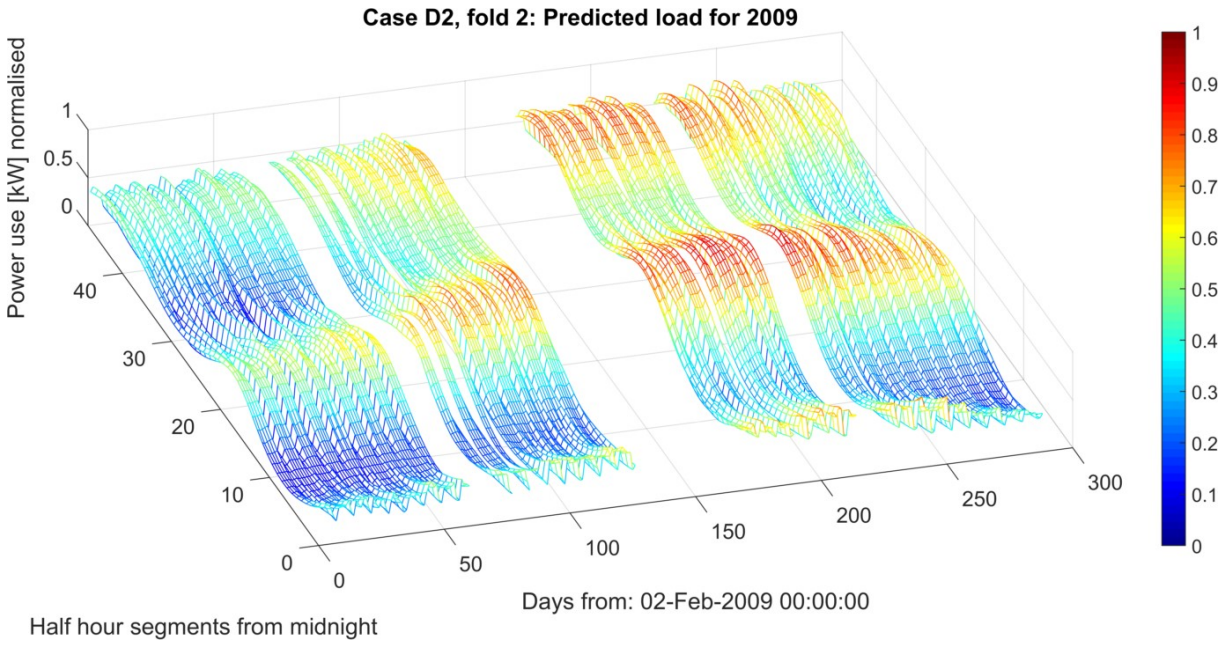

Figure A-71: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model D2.

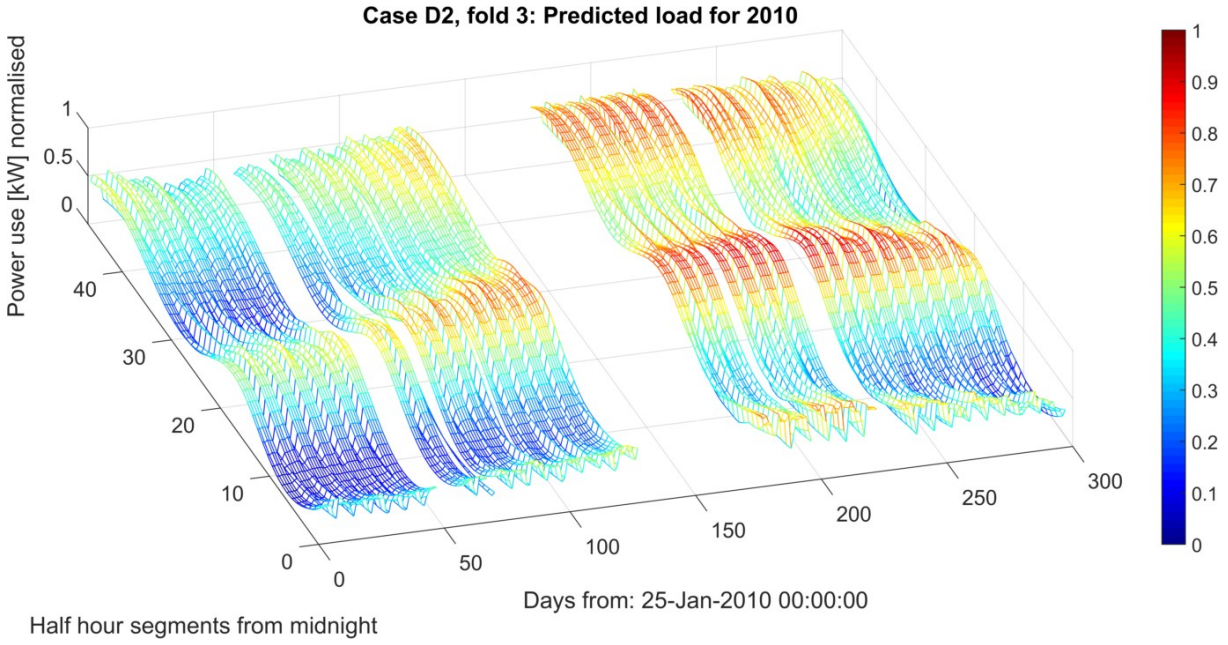

Figure A-72: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model D2.

Appendix A.13 Model D3

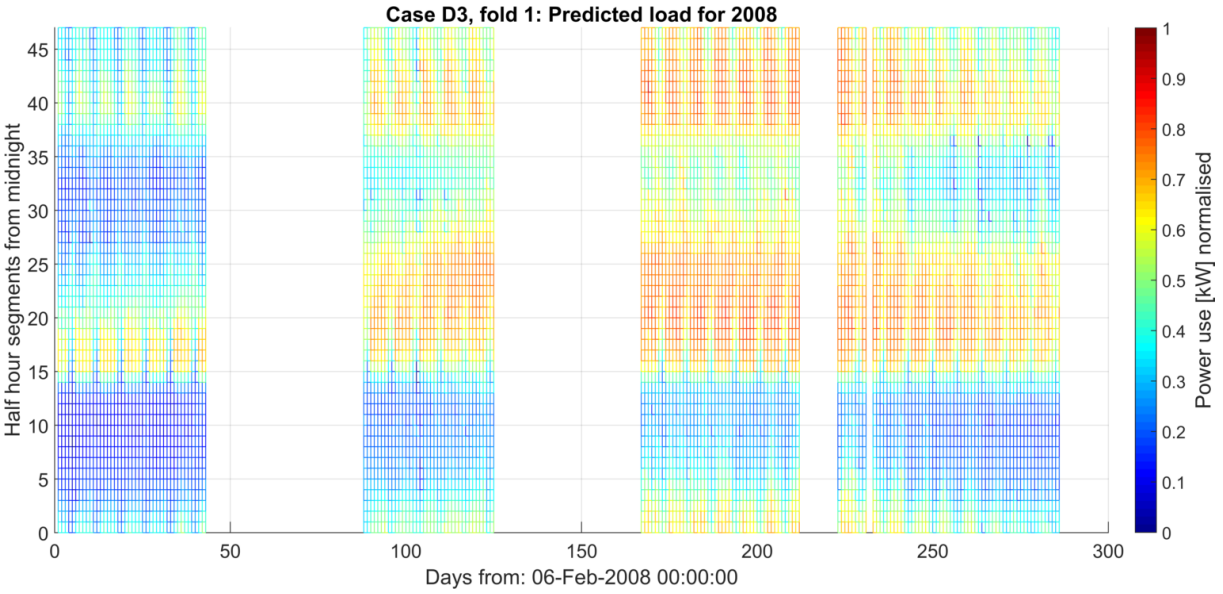

Figure A-73: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model D3.

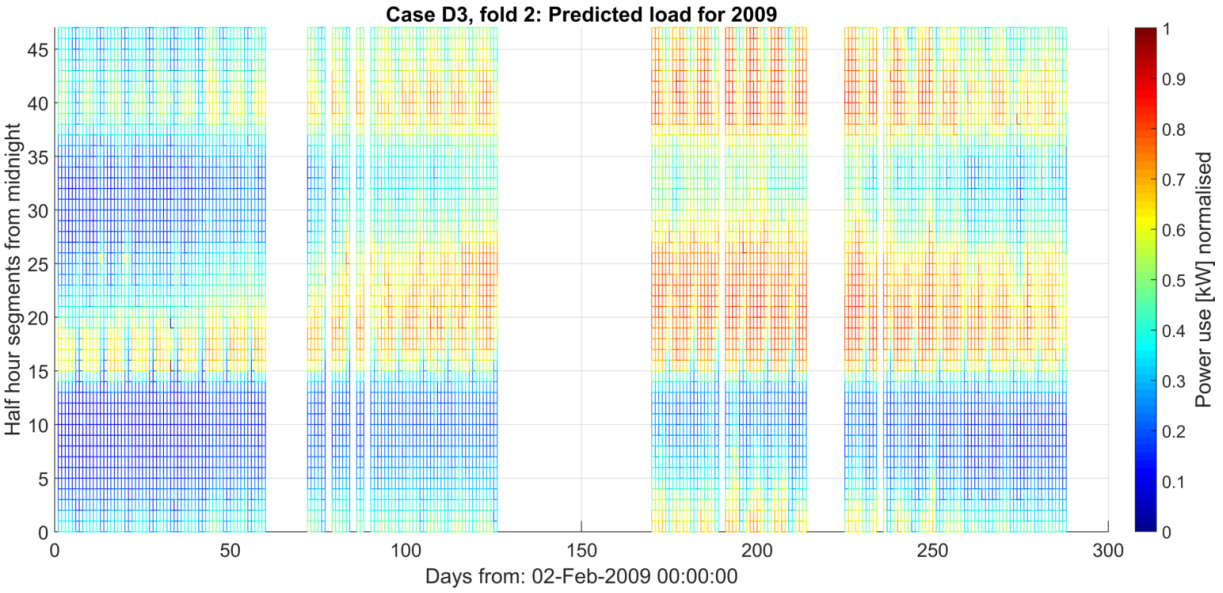

Figure A-74: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model D3.

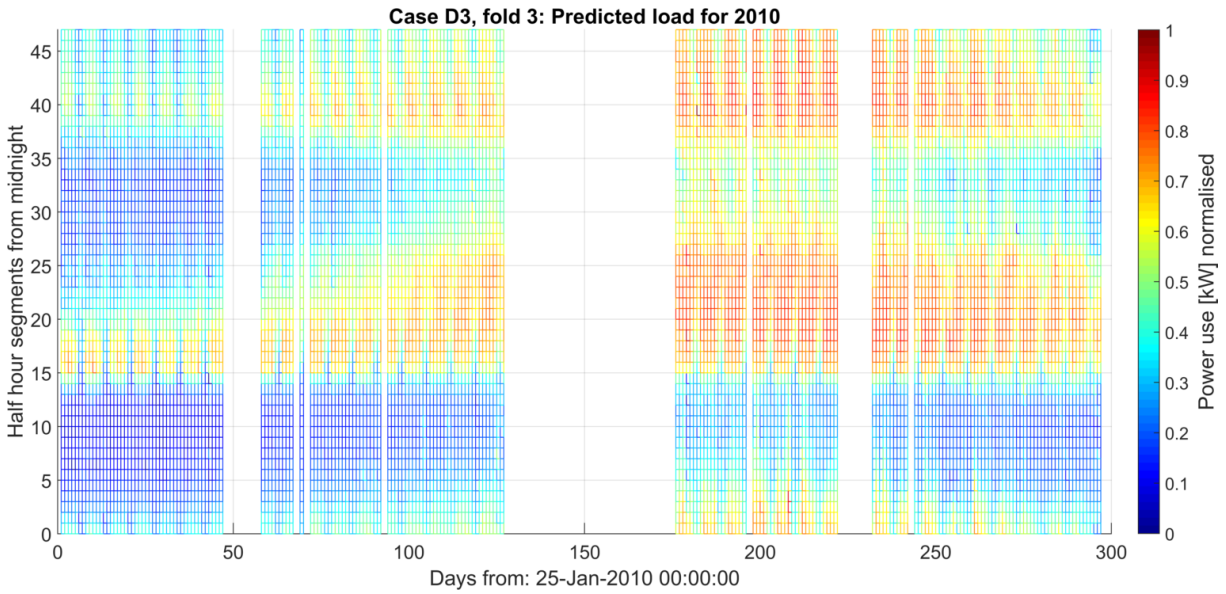

Figure A-75: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model D3.

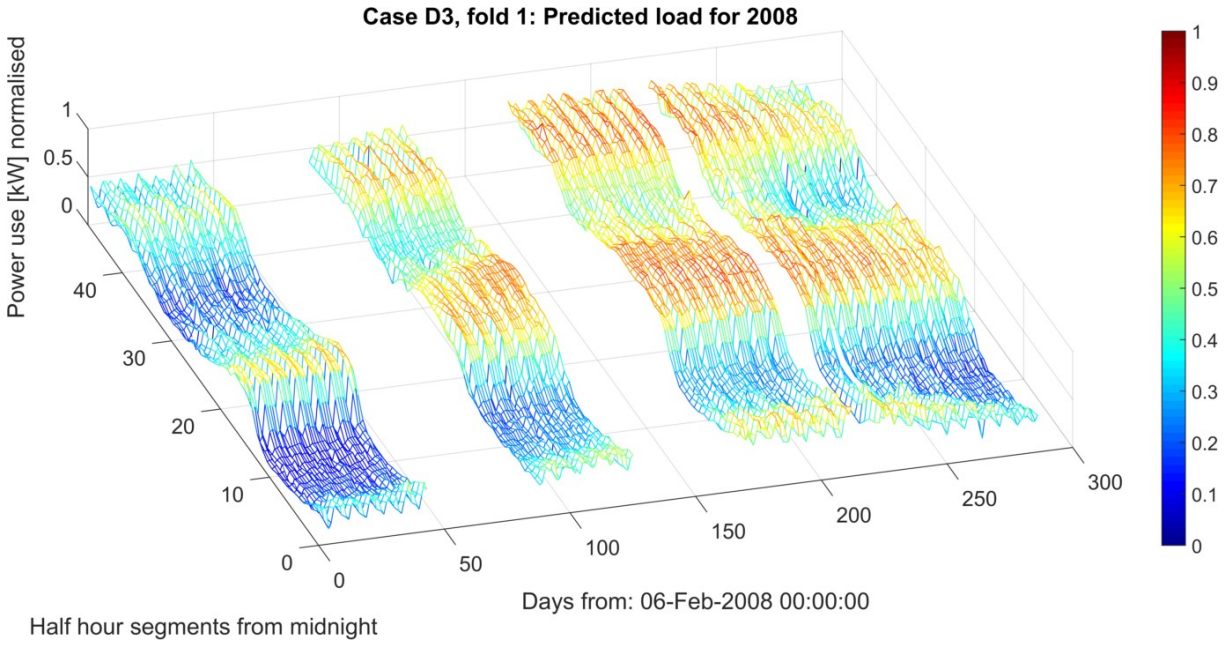

Figure A-76: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model D3.

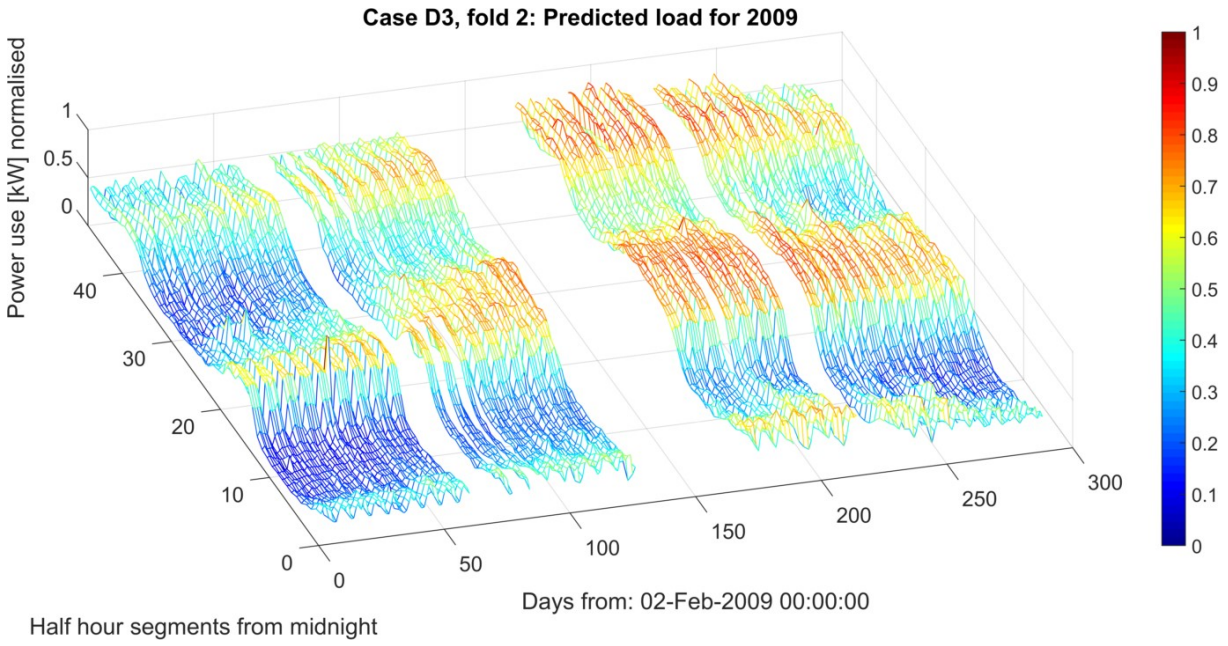

Figure A-77: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model D3.

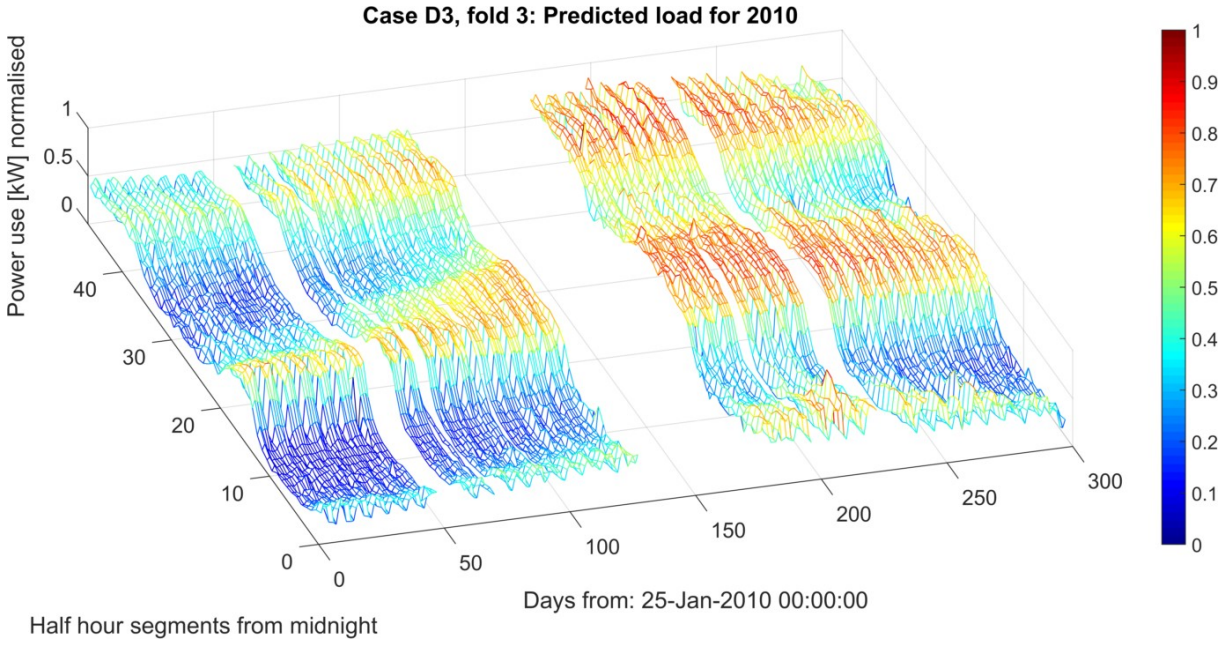

Figure A-78: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model D3.

Appendix A.14 Model D4

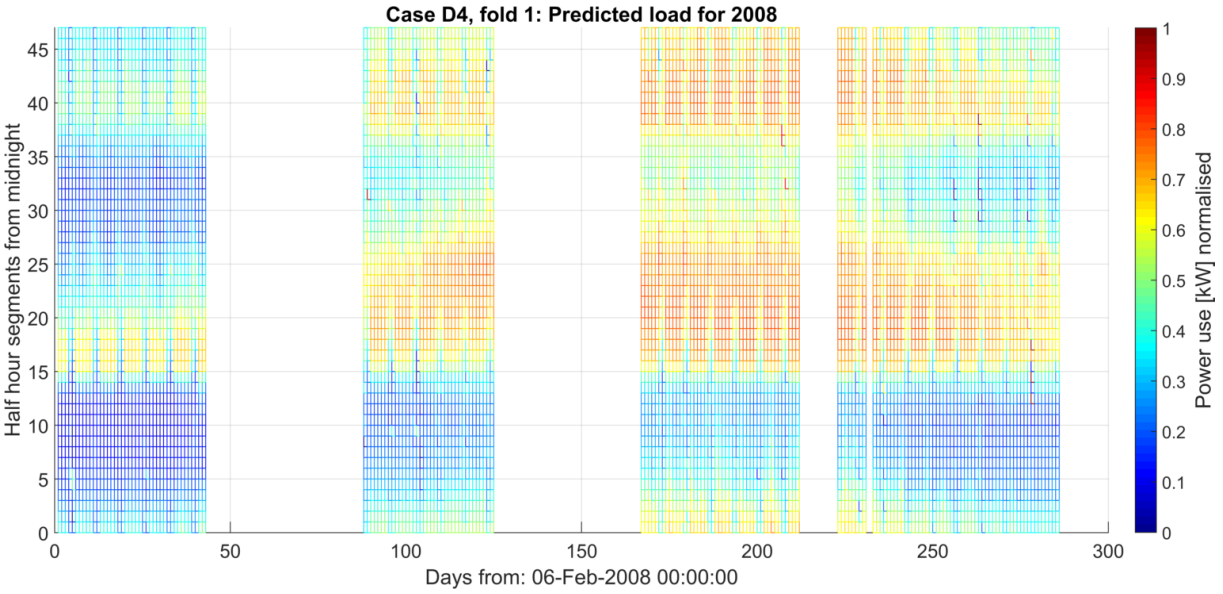

Figure A-79: Heat map of predicted load profile for 2008, fold 1 of 4, as calculated by model D4.

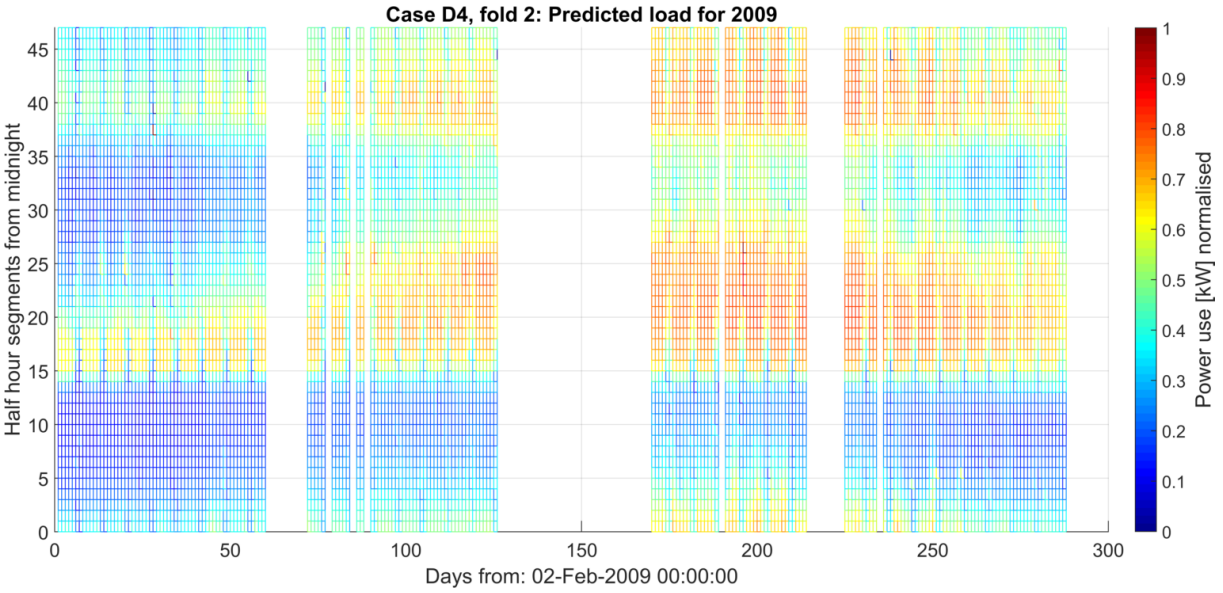

Figure A-80: Heat map of predicted load profile for 2009, fold 2 of 4, as calculated by model D4.

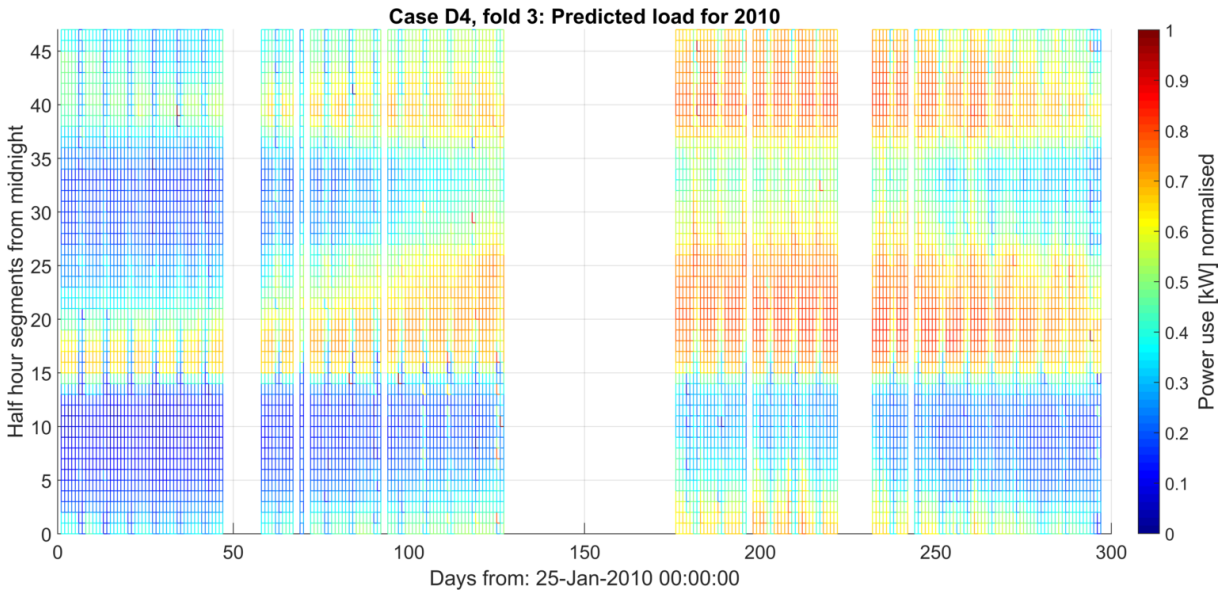

Figure A-81: Heat map of predicted load profile for 2010, fold 3 of 4, as calculated by model D4.

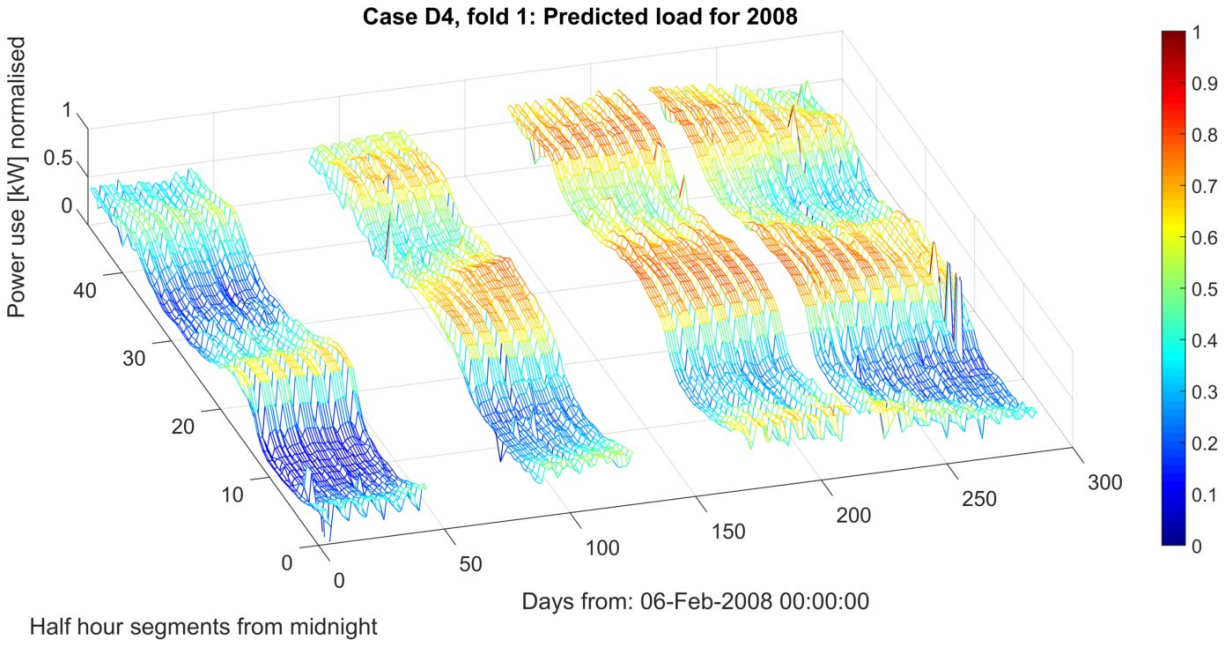

Figure A-82: 3D mesh of predicted load profile for 2008, fold 1 of 4, as calculated by model D4.

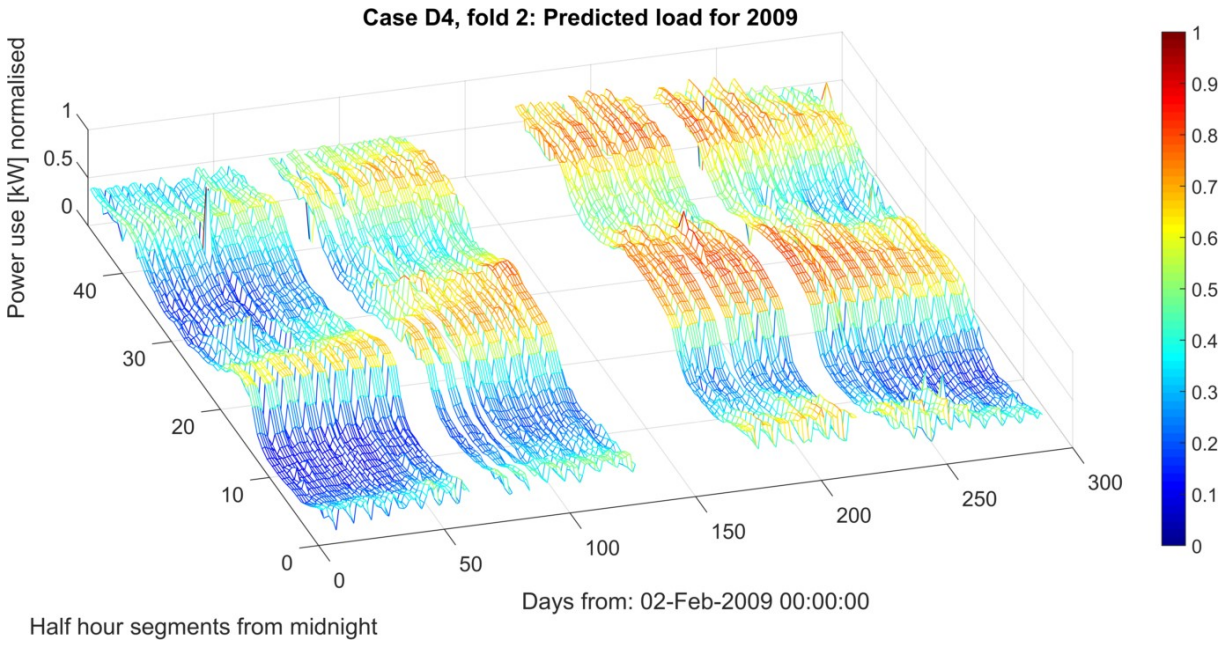

Figure A-83: 3D mesh of predicted load profile for 2009, fold 2 of 4, as calculated by model D4.

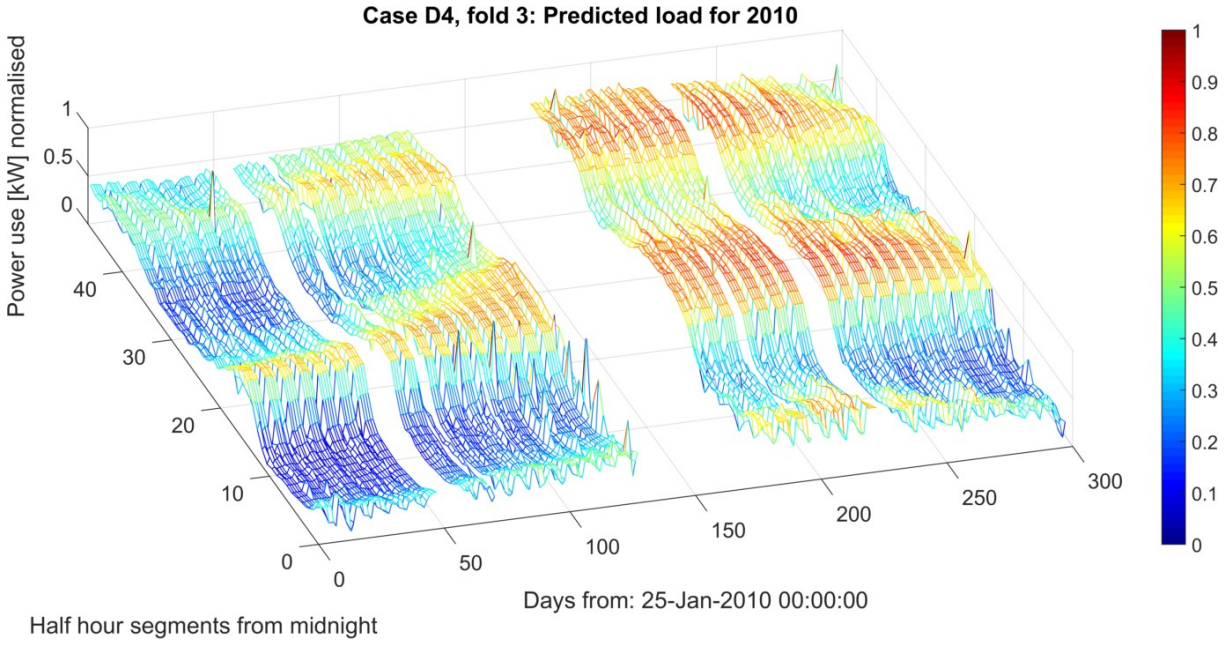

Figure A-84: 3D mesh of predicted load profile for 2010, fold 3 of 4, as calculated by model D4.

Appendix B Error Percentages

Appendix B.1 Scatterplot of the percentage errors over observed load size

Shown below are scatterplots of the percentage errors for prediction of each fold, over the size of the observed load at the error.

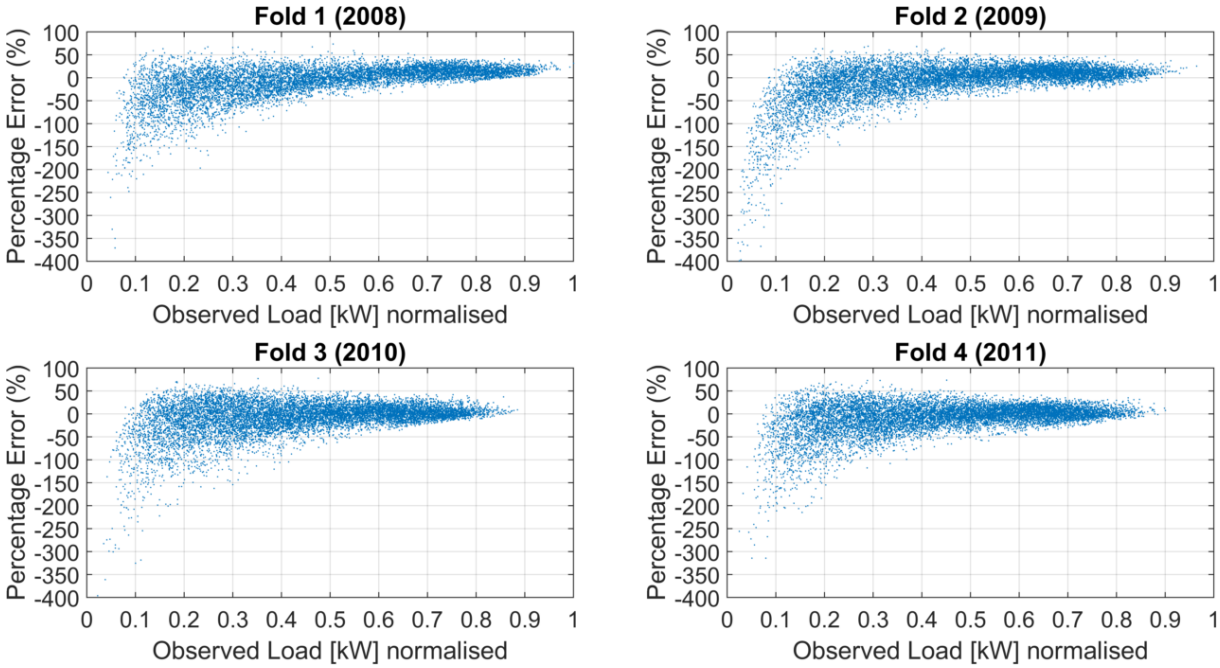

Figure B-1: Percentage error of model A2 predictions over size of observed load at error.

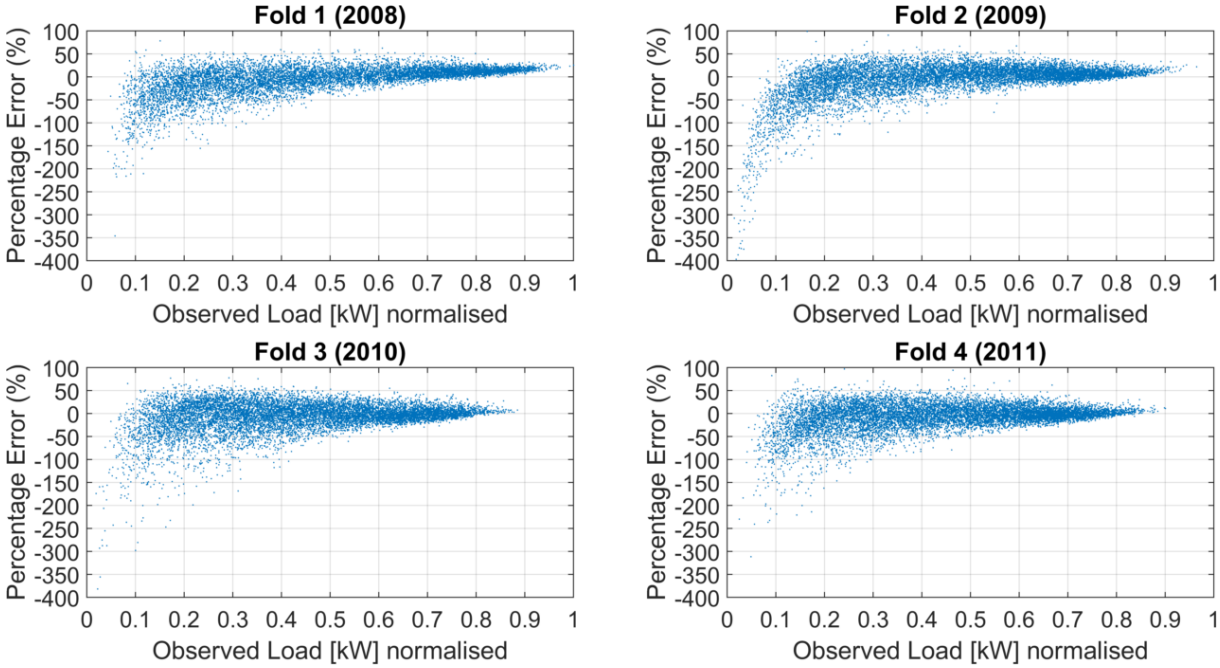

Figure B-2: Percentage error of model A3 predictions over size of observed load at error.

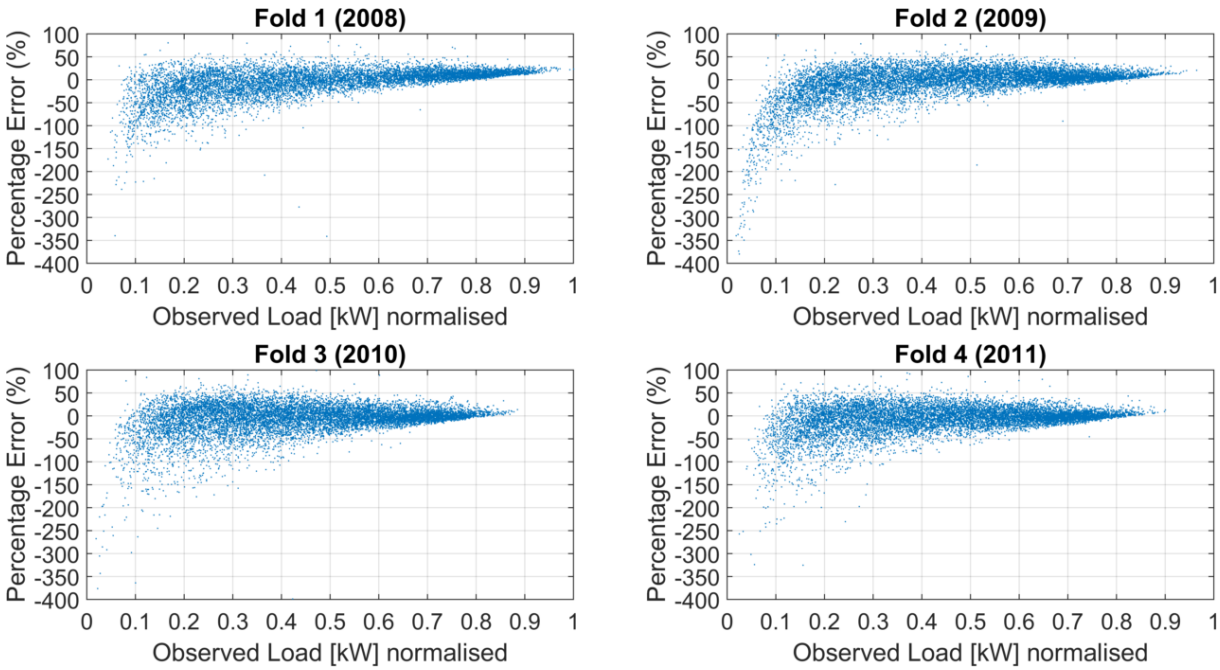

Figure B-3: Percentage error of model A4 predictions over size of observed load at error.

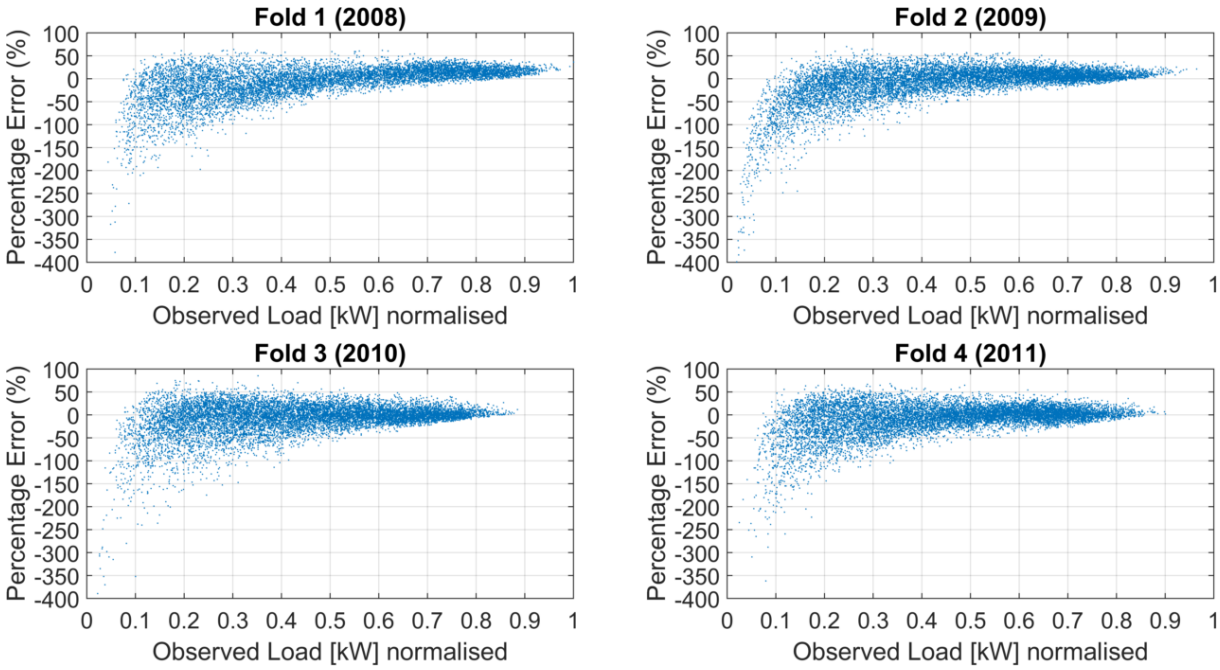

Figure B-4: Percentage error of model B1 predictions over size of observed load at error.

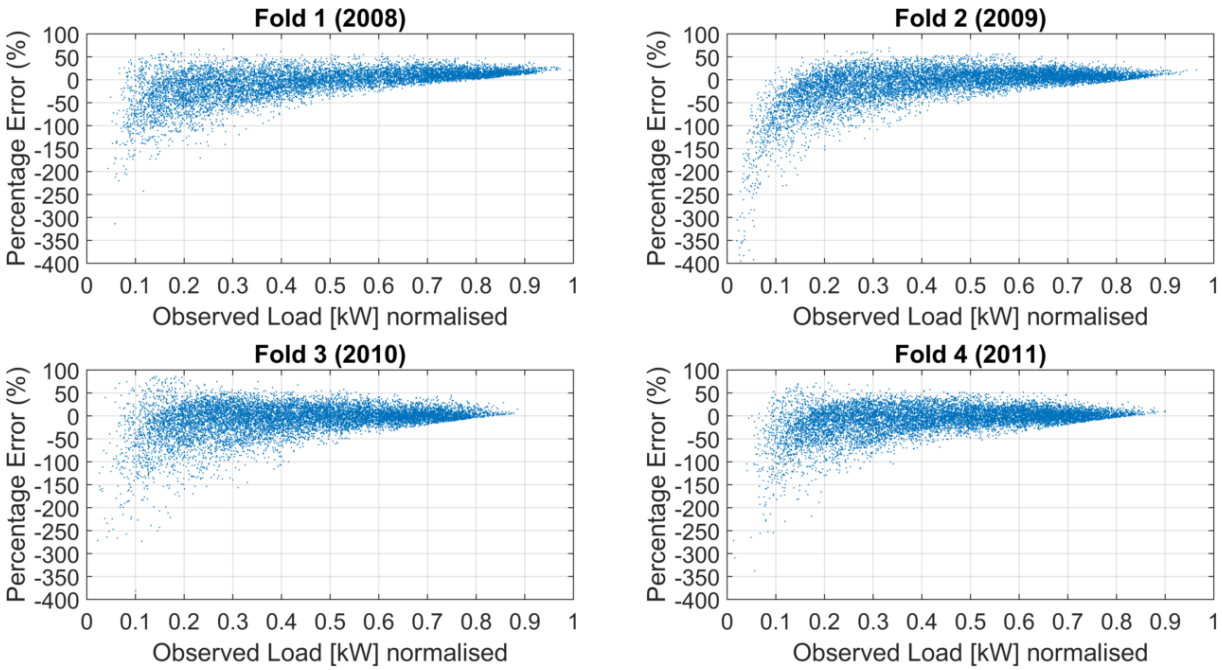

Figure B-5: Percentage error of model B2 predictions over size of observed load at error.

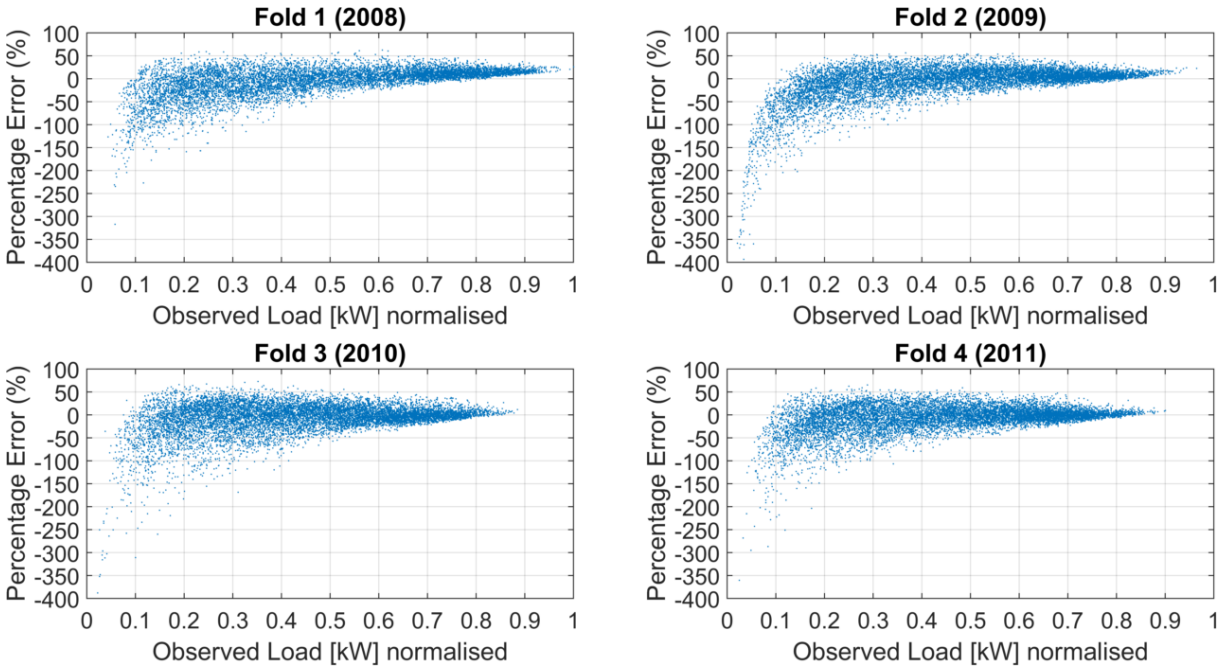

Figure B-6: Percentage error of model B3 predictions over size of observed load at error.

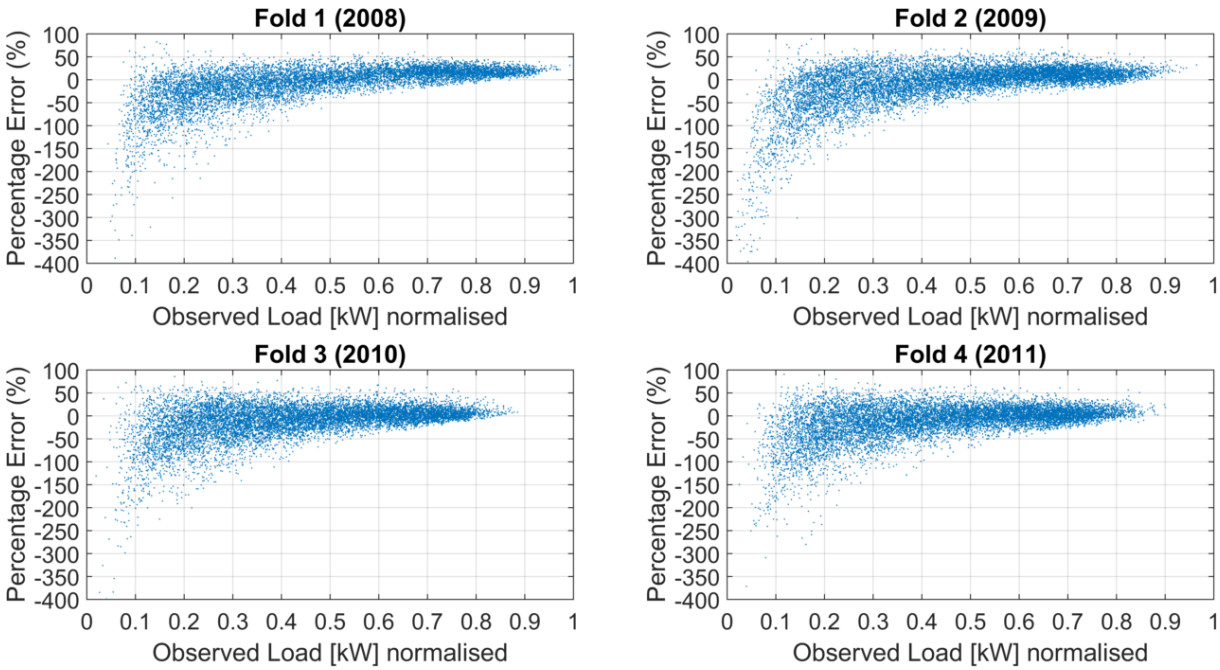

Figure B-7: Percentage error of model C1 predictions over size of observed load at error.

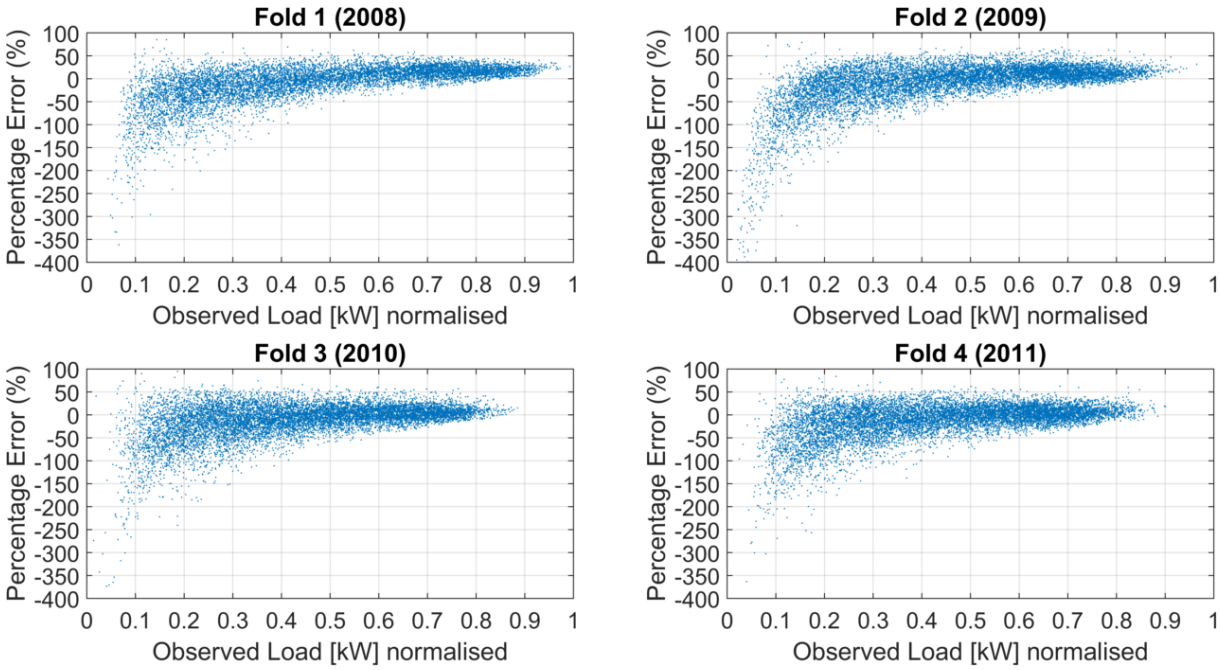

Figure B-8: Percentage error of model C2 predictions over size of observed load at error.

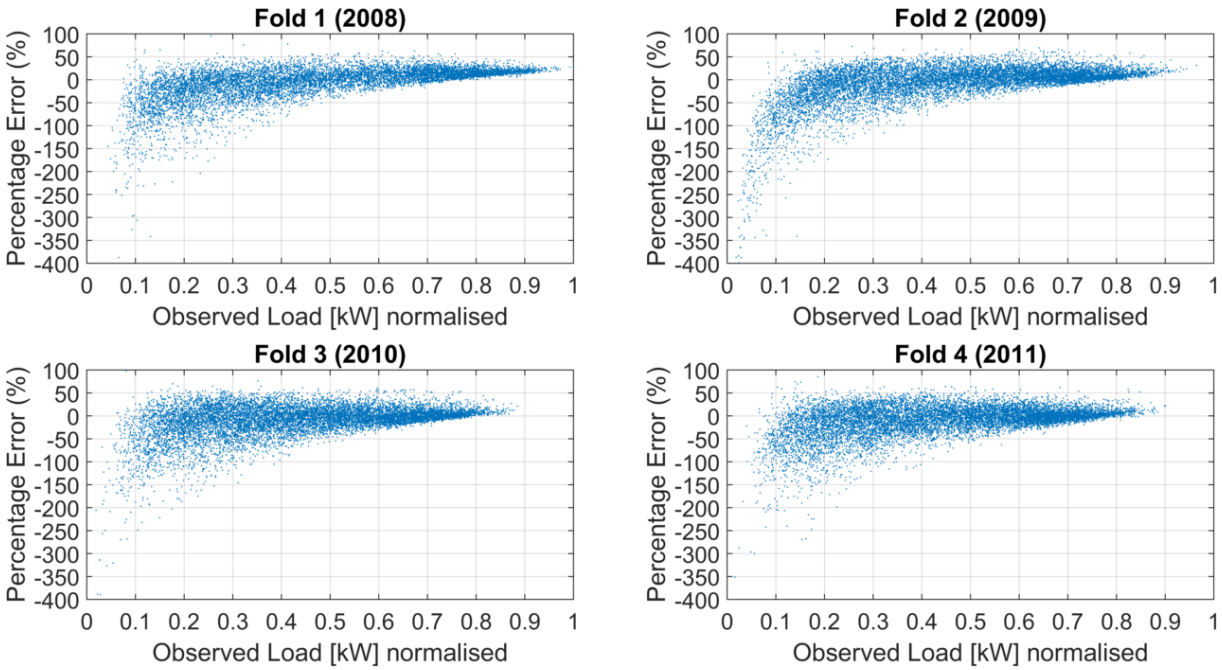

Figure B-9: Percentage error of model C3 predictions over size of observed load at error.

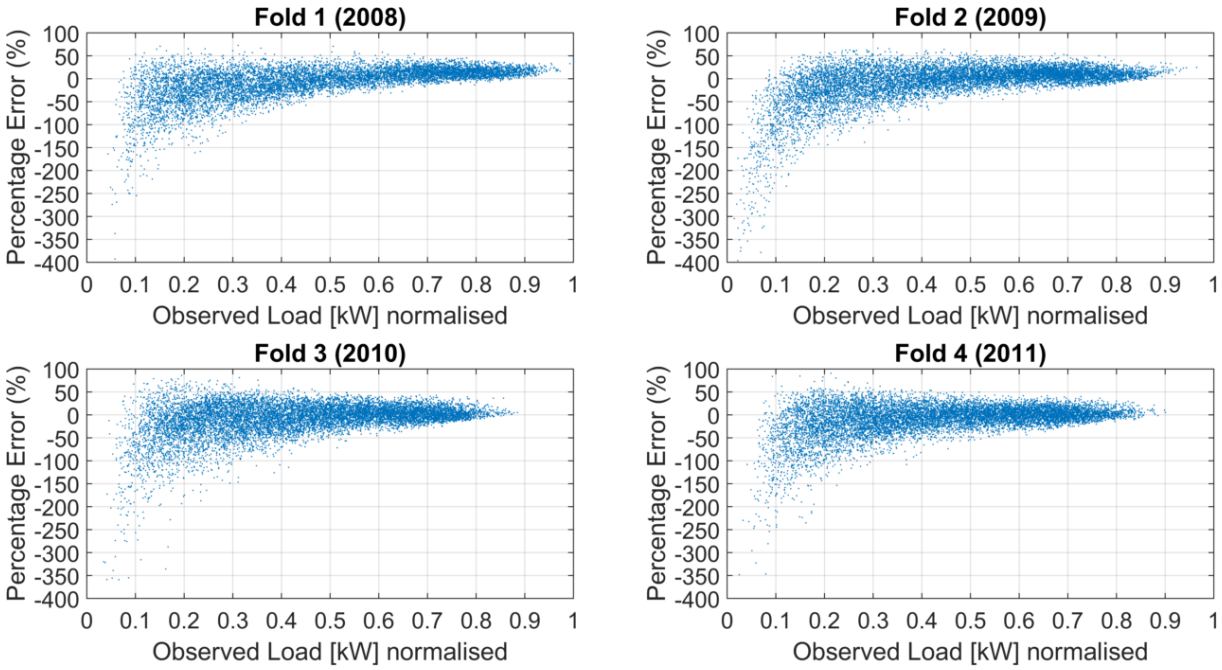

Figure B-10: Percentage error of model D1 predictions over size of observed load at error.

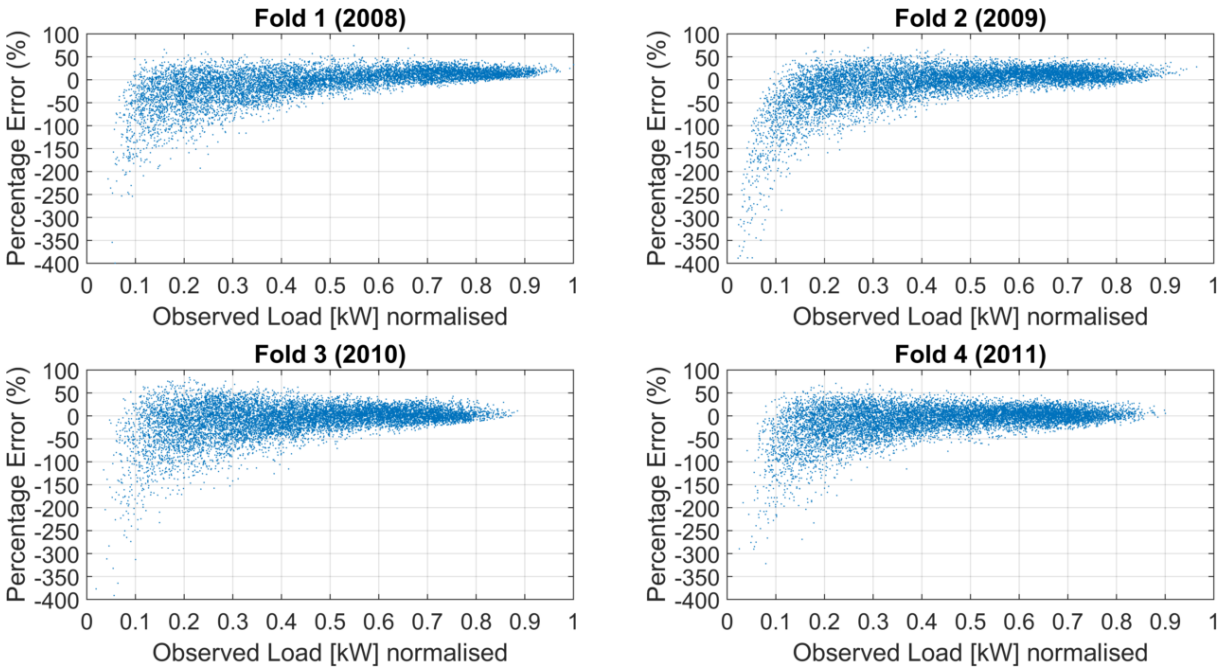

Figure B-11: Percentage error of model D2 predictions over size of observed load at error.

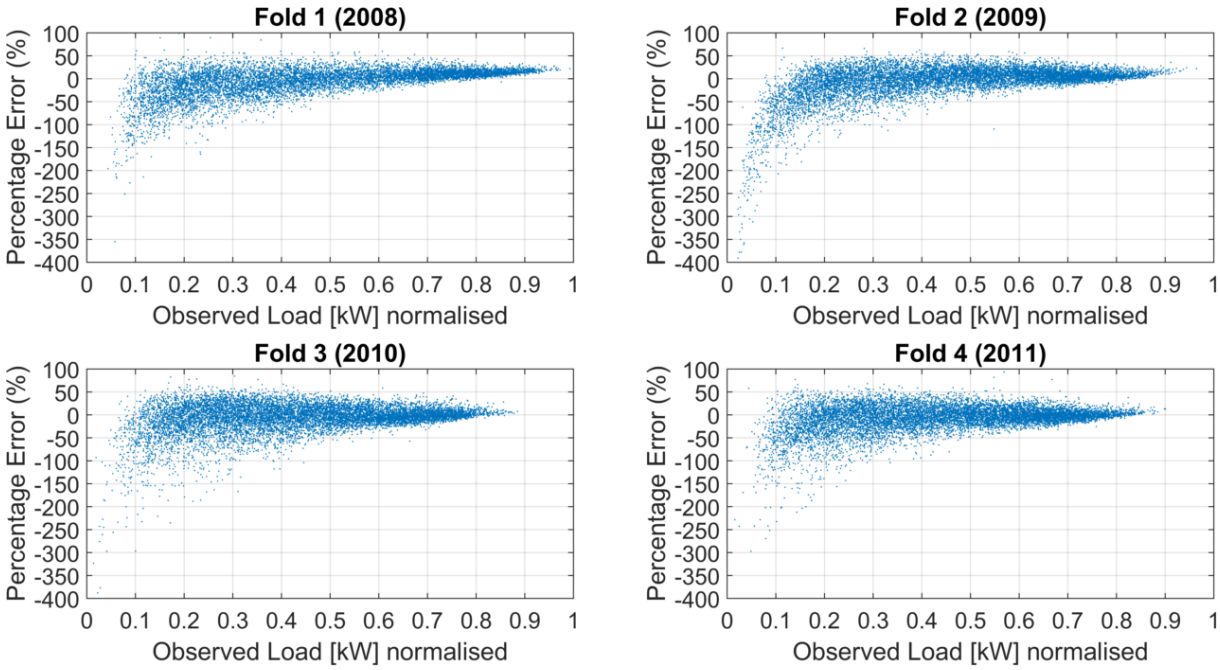

Figure B-12: Percentage error of model D3 predictions over size of observed load at error.

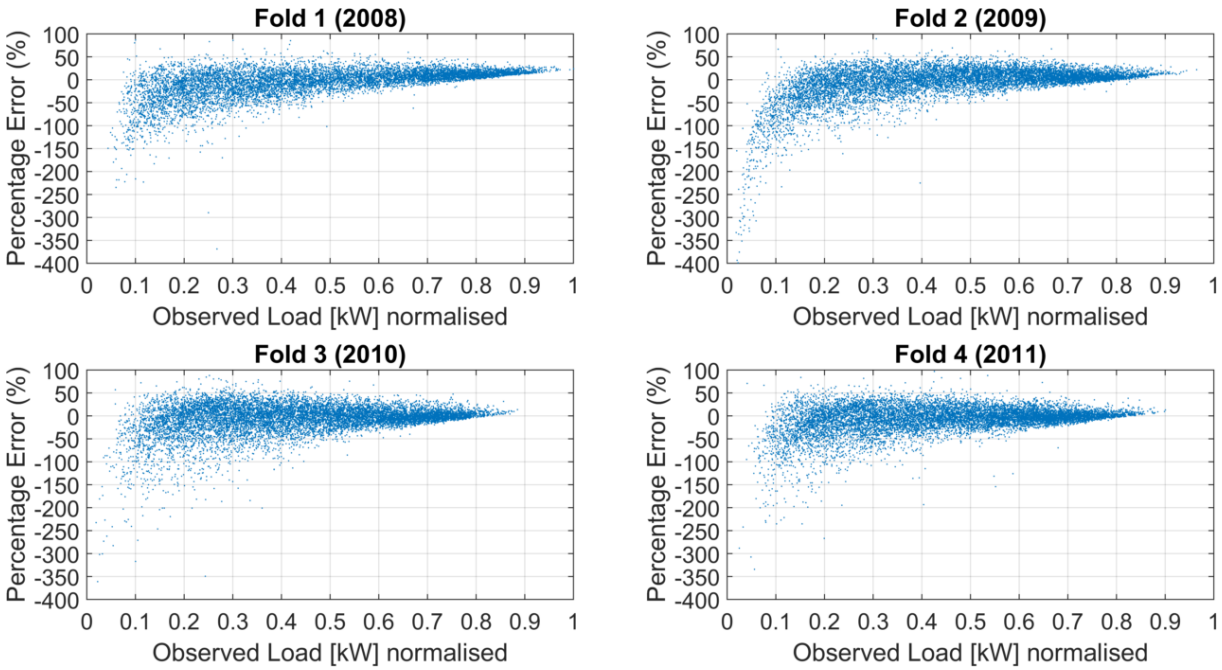

Figure B-13: Percentage error of model D4 predictions over size of observed load at error.